

Looking for a pointed role

The annual meeting of the British Association promises to be better than usual but the perennial problem of the association's proper function persists.

By the unexciting standards set in recent years, this year's annual meeting of the British Association for the Advancement of Science promises well. No doubt the venue helps — a sea-coast university is bound to give many of the participants a sense of being on holiday — but the University of Sussex has obviously been stung by its vice-chancellor's inclination to send the association packing off elsewhere into dredging into its liberal origins both to provide some interesting topics for discussion as well as a child-minding service for the benefit of participants. Fortunately, the meeting is also contained within a working week, which is convenient for those attending and which means that no bishop will be tempted to preach a sermon for the occasion, as used to be the custom. (There is no cathedral within convenient distance, in any case.) And this year the management of the British Association has taken steps that may eventually help to remedy one of the obvious defects of these occasions during recent decades — the virtual absence of up-to-date science from the usual agenda. This has been done, by persuading a number of learned and professional societies to provide symposia not merely for their own members but for those of the British Association as well. Let us hope that this and the other bright sparks in this year's programme somehow survive.

Meanwhile, the association still has to decide what its function should be. For far too long, since it ceased in the 1920s to be the kind of professional society that the nineteenth century had needed, the association has been the prisoner of uniformity of purpose. Too often, it has behaved as if its continued existence would in itself be sufficient — and then has found that the funds that kept it going (now mostly from the government under the umbrella of the Royal Society) were likely to dry up. When, in the late 1960s, the association was acutely faced with the risk of having to go out of business, the notion that other learned societies could usefully be asked to help enliven the annual meetings was widely canvassed, universally approved — and then forgotten. At that time, the association also embarked on a plan to carry out studies of public issues with important scientific connotations, but in the past fifteen years that plan has been pursued with only fitful vigour. There have been several attempts to mount public discussions of important public issues uncluttered with the rituals of the annual meetings, but these have been too few, and too uneven in quality, to point unambiguously to a new role for the association as a whole. And the admirable idea that the association might create some kind of publication that would give it a wider claim on public attention (and help to keep the wolf from the door as well) has not been diligently pursued.

Management

These are all familiar problems to those acquainted with the association. The reason they persist is also plain. The British Association is poorly managed, and the reasons for that are constitutional, not personal. Presidents vary, but even those interested in helping to get something done are there only for a year. Otherwise responsibility for policy rests with a cumbersome council, while the execution of policy is shared between an army of often autonomous volunteers and a tiny staff which is not encouraged to have ideas of its own. The results are plain in the patchiness of the programme for the annual meetings. This year, for example, there are planned both a possibly absorbing sym-

posium on the technology of satellite communications and a talk with the curious title "Biotechnology: the way ahead for botanists".

Although both the annual symposia and much of the associations other undertakings are characterized by a homespun parochialism that is no longer endearing, the association has a great deal of goodwill that it could draw on. While affection for such an old and interesting organization is no longer bankable in financial terms, it is an asset that cannot be ignored. People are willing to do what they can for the British Association (which is, indeed, how it rubs along at present). Moreover, the association has by good fortune retained among British newspapers the reputation of being a place at whose annual meetings it is possible to pick and retail to readers an account of where science stands; the fact that there tends to be little else in the way of news at this time of the year no doubt helps to ensure an uncritical appraisal of this use of resources. Then there are the people who turn up at meetings (which evidently give great pleasure) whose loyalty gives the association its annual injection of the conviction that it is doing a good job, not to mention those who are cajoled into speaking on topics of somebody else's choice, but who must usually pay their travelling expenses to do so: these and many other groups of people keep the association going, but will not do so indefinitely.

The future

To be sure about its future, the association must therefore find both a clearer concept of what it is for and a more secure means of organizing its affairs efficiently. As to its roles, there are three matters that the association should attend to, chief of which is to devise a policy. Unlike other bodies — most learned societies for example — the British Association is not prevented by its constitution from having views on questions such as the appalling arrangements for science education in Britain, the way in which research money is shared out and the suitability of British universities for the remainder of this century. To advocate that the association should develop policies for the sake of having them would of course be ridiculous, but these are all questions on which too much is decided behind closed doors, and on which independent views, publicly expressed, are in short supply. Second, the association needs to be a more effective vehicle for doing what it purports to be doing at present — providing a platform for British scientists to tell a wider public what they are up to. As budgets have been tightened and as the competition with people elsewhere has intensified, the usual reticence has given way to an eagerness to talk in public (which would serve the public interest). Third, given the association's objects and the absence of comparable informal academies elsewhere in Europe, there is a chance that the association could do a useful job if it spread its wings more widely in Western Europe. At least it would be worth a try.

The snag, it will be said, is that none of these things will be possible unless there is a stronger professional organization, which will take money, and that there is no prospect of more money coming along until the association has been able to show that it is able to perform a unique public service of some kind. But this is a recipe for getting nowhere. And the argument is false. It might suffice if the association were simply to make up its mind what kind of animal it wants to be. □



Genetic engineering

Regulation issue is resurrected by EPA

Washington

GOVERNMENT regulation of genetic engineering is once again about to become a live issue. Regulation has steadily declined as the fears that prompted the creation, seven years ago, of the National Institutes of Health (NIH) Recombinant DNA Advisory Committee (RAC) have failed to materialize. But the Environmental Protection Agency (EPA) may change all that.

In a move fraught with legal complications, EPA is asserting regulatory authority over the commercial production of genetically-engineered microorganisms — such as those designed to eat oil slicks — that would be deliberately released into the environment.

EPA's action has revived questions about the adequacy of RAC's guidelines, which are not binding on industry, about the ability to assess the risks of novel organisms and about the legal grounds for government control of genetic engineering. Under the RAC guidelines, which are mandatory only for government agencies and recipients of NIH research grants, prior permission is required for any deliberate release of genetically-engineered organisms. Industry as a whole has voluntarily agreed to adhere to the RAC guidelines, and several states and localities have passed laws making the guidelines mandatory.

Moreover, according to Geoffrey Karny of Congress's Office of Technology Assessment, tort law provides a strong incentive for industry to follow the guidelines. Should an injury claim arise in connection with a deliberately-released organism, non-negligence on the manufacturer's part would be adherence to the guidelines. "Any company would be foolish not to comply with them", he says.

RAC's continued willingness to serve as a quasi-regulatory body over industry may, however, wane as commercial applications grow. RAC's expertise has also been called into question when it comes to commercial use and to the ecological issues that pertain to environmental release. The RAC membership is largely molecular biologists and doctors. The major worry over environmental release of novel organisms is that they may find a new niche and become in-eradicateable pests, as has occurred with the introduction in the United States of such exotic plants and animals as the gypsy moth, chestnut blight and the kudzu vine. According to Karny, nearly half of the major insect pests in the United States, for example, originated abroad.

EPA has had a chance to get its feet wet in this area through its regulation of biological pest controls, which, under the Fed-

eral Insecticide, Fungicide, and Rodenticide Act (FIFRA) are considered to be pesticides; data on environmental fate and residues, performance and effects on non-target organisms are required for approval under this act. EPA officials are now working on extending risk-assessment procedures to recombinant organisms — both pest controls and others.

Outside the agency, there is considerable scepticism over the value of research aimed at developing such risk-assessment methodology. Frances Sharples of Oak Ridge National Laboratory, who has examined the problem for EPA, says that although the analogy with microorganisms such as the gypsy moth is valid, nothing in the literature provides any basis for predicting whether a novel organism will be a problem or not. EPA at present has only a handful of scientists at work on the problem, although it is said to be planning a vastly expanded research effort over the next several years.

The legal basis for EPA's regulation is only slightly less uncertain than its scientific basis. Although biological pest controls are unquestionably under EPA's regulatory authority (under FIFRA), the status of other micro-organisms, such as the oil-eating bacteria, is less clear. EPA has determined that the non-pest control organisms, as they contain altered DNA, are "new chemical substances", and thus subject to the Toxic Substances Control Act (TSCA). EPA justifies this inter-

pretation by reference to the act's legislative history, which shows a congressional intention that the law should cover all substances not already covered by the clean air, clean water or pesticide laws.

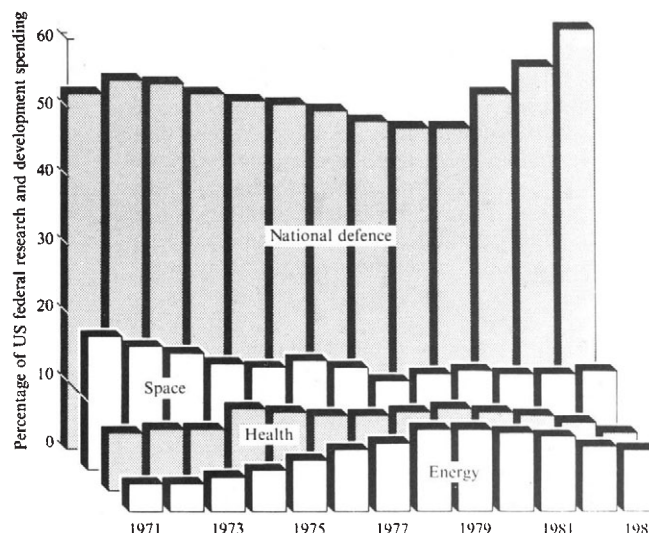
Under TSCA, a manufacturer must notify EPA 90 days before it intends to begin manufacture of a new substance and provide all available data on its safety. EPA has the choice then of requiring the manufacturer to conduct further safety tests or of regulating or even prohibiting the manufacture or use of the substance. The law does not cover research and development use of a new substance.

Karny says that while "it would be opening a real can of worms to challenge EPA" on its authority to regulate new organisms under TSCA, such a challenge "would not be off the walls" legally. Thomas McGarity, a law professor at the University of Texas who has studied the issue, says that EPA's position is "convincing, albeit risky".

Yet industry may have little to gain from such a challenge. TSCA's requirements are fairly lenient, and a challenge, Karny says "may raise questions about the companies' motives and could lead to more restrictive legislation" specifically directed at novel organisms. And, as Sharples points out, what the industry is looking for is some legal certainty from the government. "The industry is very paranoid", she says, about the legal pitfalls of going ahead with the use of manufactured organisms for such applications as enhanced oil recovery, mineral recovery or pollution clean-up. TSCA would offer the advantage to the industry of familiar procedures and some well-defined boundaries. "The lack of those boundaries is translating into inaction", she says.

Stephen Budiansky

US spending on research



PERCENTAGE federal spending on US research and development of defence, space, health and energy. The increase in energy spending after 1973 reflects the increase in non-nuclear expenditure after the oil-price rise of 1973. Source: NSF Report on Federal R&D Spending for Energy: Fiscal Years 1971-1984.

Israeli universities

Government's cuts hit hard

Rehovot

ISRAEL'S institutions of higher education, already very disturbed by a recent government decision to cut its support for universities by 6-8 per cent (in real terms) during the next academic year, are absolutely appalled by a new finance ministry proposal that this support be reduced by a further 10 per cent. Some university presidents have gone so far as to suggest that the whole system of higher education in this country may collapse if the proposal is implemented.

It is only one of many austerity measures recommended last week by finance minister Yoram Aridor in a separate attempt to forestall further deterioration of Israel's economy, and particularly a further increase in its foreign debt. During the six years that the Begin government has been in power, this debt has risen from \$11,000 million to \$21,000 million.

Professor Dan Patinkin, president of the Hebrew University and himself an economist, agrees that government spending must be reduced, but argues that support for higher education and research deserves a very high priority, and certainly should not suffer more than other sectors. He foresees "inevitable stagnation" if there are further budget cuts, which will prevent institutions of higher education from hiring younger people with innovative ideas.

Hitherto, says Professor Patinkin, Israel has been the only country of its size where both university teaching and research have been up to the "highest international standards". He warns that the maintenance of such standards will be impossible if funds are further reduced, and that "once deterioration sets in, it will be extremely difficult to reverse."

Weizmann Institute president Michael Sela hopes that some compromise will be reached, based on "extremely selective cuts." otherwise, he fears that the present university system "may not survive."

Despite past financial problems, says Professor Sela, the Weizmann Institute has managed to provide tenured positions for 50 extremely talented young scientists over the past eight years, a period during which only a handful of professors retired. This was made possible by holding down administrative expenses and foregoing any attempt to expand the institute's physical facilities. But Sela sees little chance of providing places for more tenured scientists or even of financing the work of those already on the staff if the mooted cuts are implemented.

Both Professor Sela and Professor Ephraim Kehat, vice president for academic affairs of the Haifa Technion, warn about the adverse effect of a declining university system on the country's booming science-based industries. Professor

Kehat points out that government support for the Technion, which trains a majority of Israel's engineers, has already declined by 30 per cent during the past eight years. "If it goes down much more," says Kehat, "departments will have to be closed and student intake reduced."

Professor Moshe Many, who took over as president of Tel Aviv University only last month, is already deeply involved in the struggle to prevent further reductions in government support for the university system. As a physician he described their probable effect in medical terms: "It won't be possible to deal with the cuts by putting the universities on a diet designed to eliminate body fat. They will require the amputation of whole teaching and research areas".

Professor Haim Harari, who heads the planning and grants committee of Israel's Council for Higher Education and who, with his five colleagues on the committee,

will have to decide how the cuts are to be implemented, has yet to comment on the situation (because he is overseas). But his introduction to the planning and grants committee's 1981/82 report, published this May, shows his probable reaction. In that statement he warned that the cutbacks already implemented had resulted in a lowering of standards. For example, while the student population at Israel's universities has increased by 30 per cent during the previous decade, their academic staff has declined by 3 per cent. There was also less money for equipment, libraries and the hiring of younger researchers. Unless this situation improved, he wrote, "Israel's security, economy and culture will suffer in the not too distant future".

Will protests and warnings from the university persuade the government to abandon the 10 per cent cut? No one knows. But at a time when the finance minister has also recommended that pensions be taxed, maternity grants ended and hundreds of high school teaching positions eliminated, it will not be an easy struggle.

Nechemia Meyers

US nuclear industry

Glimmers of hope

Washington

THE American nuclear power industry, which has not received a single domestic order for a new reactor since 1978 and suffered 18 cancellations last year, may get a helping hand from the State Department in finding new customers overseas. After talks with other government agencies and the principal manufacturers, the department is hoping to create a Nuclear Safety Training Academy that could offer cut-price advice and training to foreign countries embarking on civil nuclear power programmes.

Although it has fallen on hard times, the American nuclear industry remains the biggest in the world and enjoys the warm support of the Reagan Administration. The academy would be a sign of the United States' commitment to nuclear safety and give American firms a market edge by familiarizing potential customers with US technology and safety practices.

As envisaged, however, the function of the new academy would be limited. Several firms involved in talks with the State Department were lukewarm about a proposal that might duplicate the extensive training programmes already offered by the major manufacturers, universities and national laboratories. The General Electric Company, for example, offers 82 courses in nine different areas of nuclear training and maintains simulators for its boiling water reactors. Training services are part of the companies' sales packages.

Instead, the new academy would consist of a small foundation, jointly funded by government and industry, whose main job would be to coordinate existing training

courses. A preliminary report on the academy's work says industry would be "intimately" involved in its activities and individuals attending the academy would spend much of their time with utility companies, vendors and architect-engineers.

Supporters of the academy believe it could nevertheless fill an important gap in safety training, by attracting high-level utility managers and policy-makers as well as plant managers and operators. Safety courses run by plant manufacturers are usually oriented towards lower-level operators. The academy would design a special course for senior officials in government and industry, lasting several weeks and including visits to the Nuclear Regulatory Commission, the Department of Energy, national laboratories and utilities at various stages of construction.

Initial plans for the academy have been approved in principle by a working group of government, industry and research representatives. Several knotty problems remain to be sorted out, however. A "balanced" curriculum acceptable to the Nuclear Regulatory Commission as well as the manufacturers must be devised. And a person of independent standing, capable of commanding international respect, must be persuaded to become its director.

Meanwhile, in a separate initiative, the International Nuclear Studies Group is meeting in Geneva this week to consider establishing an International Commission on Nuclear Safety. Modelled on the International Commission on Radiological Protection, the commission would be asked to set universal standards for plant safety.

Peter David

US space programme

NASA plans a space station. . .

Washington

THE National Aeronautics and Space Administration (NASA) is urging the United States to take its next historic step in space — the construction of a permanent manned space station that could become the centre of a space-based manufacturing industry as well as the jumping-off point for manned trips to the Moon and Mars.

Under the guise of a "modest" planning exercise, NASA has lobbied assiduously to make the space station the centrepiece of its activities to the end of the century and beyond. The agency has told the White House that a space station is the logical sequel to the shuttle and offers the best prospects for maintaining American leadership in space. It seems likely that President Reagan will agree.

Space stations have come and gone from NASA drawing boards for more than two decades. In the 1960s, plans for a space station were pushed aside by the Apollo Moon programme and in 1970 lack of political support forced NASA to drop the concept of a space station in favour of the shuttle. The agency's administrator, James Beggs, now wants an all-out effort to launch the first elements of a space station by 1991, which he estimates could cost less than £9,000 million.

Until recently, much of NASA's planning effort has been devoted to finding potential users for the space station, but in recent months agency officials have become less coy about the technical details. What has emerged is not a single large space facility but a network of manned and unmanned platforms, laboratories and satellites linked by a new generation of orbital transfer vehicles.

The vanguard of the space station in 1991 would consist of a main base at an orbital inclination of 28.5°. With a crew of six to eight, the base would include two or three pressurized modules for research and development, with a volume of about 120 cubic metres. A separate unmanned platform would be placed in polar orbit. By the end of the century, the space station would be expanded to provide space for a crew of 12 to 18 and the polar station might be turned into a manned facility.

These plans are so tentative, however, that NASA only recently lifted a self-denying ordinance banning the publication of drawings of the proposed station. A NASA official told Congress last week that at the end of meetings with foreign countries interested in the space station, all doodles had been carefully destroyed.

The ostensible reason for NASA's reluctance to define its plans is that it wants to be sure that it has the best possible design before it is committed to specific pieces of hardware. Another reason, not publicly stated, is that the agency expects to find it easier to win political friends for the space

station while the station's capabilities and uses are still somewhat vague.

The Department of Defense (DOD), for example, has so far refused to enthuse about the military uses of a space station. Richard DeLauer, Under Secretary for Defense for Research and Engineering, told a recent NASA symposium that the Pentagon had been unable to identify a single military function that could be done better by a space station than by an unmanned spacecraft. DOD is nevertheless deeply involved in hypothetical discussions about the station's design, enabling NASA to argue that a space station could make great contributions to national security.

According to Beggs, the space station could ultimately evolve into a command post or operations centre for DOD, as well as a storage facility and a base from which military satellites could be serviced. Later, Beggs adds, there would "probably" be two stations, one in polar orbit primarily for DOD use and one in equatorial orbit (the 28.5° main base) primarily for NASA.

NASA is using a similar manoeuvre to overcome the scepticism of the US space science community. The National Academy of Science's space science board has been drawn into the space station planning process and asked to consider how space scientists could exploit a space station if one were to be built. The board has not, however, been asked (by NASA) whether it believes there is a scientific justification for investing in a space station in the first place. If it were to be asked, the answer would almost certainly be no. Although it has so far observed a prudent

public silence, the board has been working on a policy statement spelling out its belief that money for space science could be better spent. The draft statement complains that current budget levels support only a small fraction of existing mission possibilities for space science and that "few" disciplines in the life sciences and none in the physical sciences need a manned space station at present. In the long run, the statement concludes, a space station could provide a "significant opportunity" for some disciplines, but space science is still just beginning to learn how to use the capabilities of the shuttle.

The National Academy of Science's space applications board has been less hostile, but also fears that by committing itself to an enormous engineering project like the space station, NASA will be forced to divert money from applications to pay for the inevitable cost overruns in development.

Despite the scepticism of some groups, Beggs believes that the White House will give the project its blessing within the next 6 to 12 months. The senior interagency group responsible for space policy is expected to receive a recommendation from its space station working group by November. Chaired by NASA, the space station group is likely to argue that without a single big engineering project like the station, NASA will be unable to maintain its preeminence in space. The Soviet Union, NASA has been pointing out, has already developed an entirely automated resupply unit for its Soyuz 7 vehicle, and recently attached a propulsion unit that will be able to move the craft into new orbits.

It will not be entirely plain sailing, however. The president's science adviser, George Keyworth, has maintained a

. . . Soviet union almost there?

In pressing for a space station, the US space agency is following the path taken by the Soviet Union. The Soviet space programme has long been committed to the idea of establishing a permanently staffed orbital station, and in 1976, a leading space-planner, Academician Georgii Petrov, talked about a staff of 100 people, with some twenty or thirty cosmonauts on board at any one time.

The new Soviet space "tug" launched on 27 June as *Cosmos-1443*, and now operating as part of a manned orbital complex with *Salyut-7* and *Soyuz-T*, marked an important step in the Soviet programme.

The new craft, with a length of 13 m, maximum diameter more than 4 m and mass (including return module) of 20 tonnes, is the largest cargo spacecraft ever launched to rendezvous with a *Salyut* station. It has more than 2.5 times the cargo space of its forerunner, *Progress*, and, unlike *Progress*, has its own power source — a solar cell of area 40 m². The return module allows up to 500 kg of

"useful cargo" — photographic film, semiconductor materials produced in space, equipment for reuse in ground-level control experiments — to be sent back.

The early recovery of film will be of particular importance to the Earth resources survey, which, a *Pravda* article stressed last week, makes a major contribution to the "practical character" of the manned orbital programme. Any suggestion that it could also be useful for military reconnaissance would, of course, be hotly denied by the Soviets who consistently contrast the "peaceful" aims of their space programme with the "military" aspects of the US space shuttle.

Not everyone abroad, however, accepts the Soviet assurances. The *Cosmos-1267* craft, used to supply and enlarge the *Salyut-6* station, and subsequently used to effect a re-entry burn when the complex had finished its useful life, was believed by a number of US analysts to have been armed with anti-satellite homing devices.

Vera Rich

carefully ambiguous public attitude. In a delphic speech in Seattle recently, Keyworth hinted that it would be a mistake to plunge into a space station project without spelling out precisely what the next step in space would be. Significantly, he dismissed the argument that Soviet achievements with Soyuz had stolen a march on the United States.

The two most significant deciding factors will be the White House's perception of the military use of the space station and the significance it assigns to private sector interest in space-based ventures. DOD's lukewarm position on the space station will be taken with more than a pinch of salt: the Pentagon learned with the shuttle that by appearing to be unimpressed it could gain access to an attractive space asset without having to pay for its development out of its own budget.

As for the private sector, NASA has spared no effort to persuade industry that a space station could lead to real manufacturing possibilities. The private management consultants, Booz, Allen and Hamilton, said last week that under a contract from NASA they had determined that there was enough economic potential in space to persuade a significant number of companies to invest in the space station. The most promising opportunities were in biological processing, production of high-performance catalysts, a fee-for-service laboratory in orbit and the production of high purity iridium crucibles for making gallium arsenide.

Peter Wright Wood, Booz Allen's senior vice president, warned, however, that several important obstacles had to be surmounted before many companies would invest in the project. Most companies complained that NASA's commercial procedures were too cumbersome and insisted that some way must be found to protect intellectual property that would be acquired by an investment in space.

NASA has also begun extensive talks with foreign governments interested in becoming users of a space station. Canada, France, Germany and Japan, as well as the European Space Agency, have carried out their own studies of missions that could be mounted from NASA's space station. Until the project is formally approved, however, NASA will not talk to foreign governments about collaboration in the station's planning and construction.

Even if the White House agrees to approve the space station, NASA will still have to deal with the sceptics in Congress. Some committees have already indicated that an unmanned space station using refinements of existing hardware offers a cheap way to achieve many of NASA's goals. But most congressmen are enthusiastic. In a recent survey by the Political Action Committee for space 26 senators and representatives favoured a space station, 3 expressed opposition and 13 were undecided.

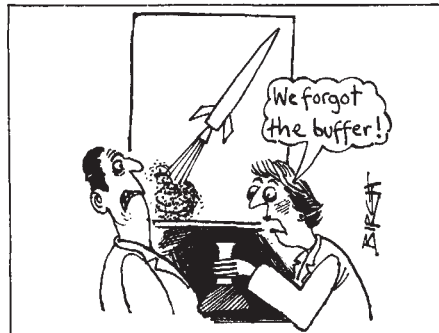
Peter David

Cell astrobiology

Shuttle separation of islet cells

St Louis

ON the eighth space shuttle flight, due to be launched on 30 August, the McDonnell Douglas Corporation will be using continuous-flow electrophoresis to try to separate insulin-producing beta cells from the other cells of the pancreas. The procedure could have potential as a way of obtaining cells for transplantation into human diabetics. On board the shuttle *Challenger* will be dog pancreatic cells provided by Drs David Scharp and Paul Lacy,



researchers at Washington University Medical School, St Louis. They are interested in separating pure beta cells for transplantation experiments from the 98 per cent of other cells in the pancreas.

Whether they will need zero gravity in space to get the purity at which they aim is not yet known, but the absence of gravity allows electrophoresis to be done without the usual Earth-bound constraints. In space, high voltages can be used without setting up convection currents that disrupt the separation, and the concentration of the sample is limited only by the solubility

of the sample, not by the density of the buffer used.

Many questions remain to be answered about how to handle the cells before, during and after the flight — such as what kind of buffers to use and whether to use fresh or frozen tissue. This first attempt at separating beta cells in space will be used to work out the handling and separation procedures. The cells harvested from the continuous-flow electrophoresis will be tested for viability on their return.

McDonnell Douglas started working in 1977 on continuous-flow electrophoresis for separating materials in space. It has joined with the Ortho Pharmaceutical Corporation to use the procedure for drug purification.

The company plans three phases of work in space. First, it will perform separations on board the shuttle. Then it plans to launch free-flying satellites that contain automatic equipment for the separations. These satellites will be serviced by the shuttle, which will be used to load raw materials and recover finished products. By the turn of the century, McDonnell Douglas hopes to use the facilities of an orbiting space station for the work.

Much remains to be done before purified human pancreatic beta cells could be transplanted into diabetics. Even if purifying the required number of cells turns out to be feasible, the details of transplantation remain to be worked out. In any case, a procedure that works for purifying beta cells could probably be adapted to other cell types that are difficult to purify.

Karen Freeman

Yellow rain

More toxin victims claimed

Washington

MORE evidence in support of the US State Department's contention that toxin weapons are being used in South-East Asia was forwarded to the United Nations last week, bringing to 20 the number of supposed victims found to have fungal toxins in their blood. The trichothecene mycotoxins T-2 (which the State Department says is the major toxin used in "yellow rain") and HT-2 (a metabolite of T-2) were detected in the blood of four subjects, 5 days to 2½ months after allegedly being attacked on different occasions between November 1981 and March 1983 in Laos and Kampuchea. By contrast, no toxins were detected in five control subjects of similar background who did not claim to have been attacked. While the use of control samples in this latest set of data appears to buttress the State Department's claims, no explanation is offered for the failure to detect toxins in at least seven other subjects who reported

being victims of three of the four attacks.

At a briefing last week for reporters, Robert Dean of the State Department dismissed criticism that focuses on the scientific data, saying that sceptics such as Matthew Meselson of Harvard are "working from a very limited body of data" that does not include intelligence information which "has really convinced us beyond a shadow of a doubt" of the Soviet use of toxin weapons in the region. The State Department, citing security reasons, has refused to make public any of this intelligence information.

One bizarre note was interjected into the latest scientific data by the State Department's assertion that the 6 November 1981 victim, described as a Lao resistance fighter, had been attacked with a "toxic agent grenade". The victim was the only witness to the incident. Previous reports have referred to aircraft-delivered sprays or artillery shells.

Stephen Budiansky

Neutrino search

New plan causes a stir

Washington

CONSTRUCTION of the prototype for a massive neutrino detector is to begin soon on the Pacific Ocean floor off the coast of Hawaii. Unlike existing neutrino detectors, "Dumand" (for "deep underwater muon and neutrino detection") will search for very high energy neutrinos thought to be produced by astronomical sources such as supernovae, and will take the unconventional approach of using an enormous mass of ocean water as the primary detector.

The Department of Energy (DoE), which has funded preliminary studies of the concept for several years, has approved construction of a small prototype to be completed in late 1984. Ultimately, a series of 500-metre-long strings of photo-detectors will be deployed that will pick up the Cerenkov light produced by muons passing through the water above, the muons having been generated in turn from the high-energy neutrinos as they undergo interactions in the water. The completed device will contain a total of 756 detectors, enabling the direction of the neutrino source to be determined.

In addition to the University of Hawaii, which is serving as the centre for the project, six other American universities are participating along with institutions in Japan, West Germany, and Switzerland. Total construction costs are estimated to be \$12 million, about half of which is expected

to come from DoE.

DoE's support of Dumand is causing a certain consternation among some neutrino physicists in light of DoE's refusal to support a solar neutrino detector that Brookhaven National Laboratory has proposed to build. This detector, which was strongly endorsed by the Field commission study on priorities in nuclear physics, would contain 45 tons of gallium that could pick up very low-energy solar neutrinos, and thus answer the riddle of why the neutrino flux measured by Brookhaven's existing detector (which uses chlorine) is only one-quarter that predicted by current solar models. The gallium detector, which could pick up the basic neutrinos produced in the fundamental solar reaction of $p + p \rightarrow D$, could determine whether the solar models are wrong or whether the neutrinos travelling from the Sun are somehow altered en route.

DoE, while agreeing to support the operating costs for the experiment, has refused to spend anything on acquiring the gallium, which costs some \$600,000 per ton — about \$25 million in all. The Max Planck Institute has offered to supply twenty five per cent of the gallium; Brookhaven officials are reportedly preparing a proposal to the National Science Foundation for support to purchase the remaining seventy five per cent.

Stephen Budiansky

Failed Soviet collaboration

THE Dumand project was originally conceived as a joint US-Soviet venture. In September 1979, three months before events in Afghanistan cast a cloud over such collaborations, Academician Moisei Markov, head of the Soviet side of the project, spoke warmly of the "joint theoretical elaboration and practical implementation" which would put into practice an idea mooted by the Soviet side at the end of the 1950s, the detection of high energy neutrinos under water of ocean depths.

The Soviet intervention in Afghanistan spelt the end of all formal Soviet participation in Dumand, although individual American participants admit that they have received valuable theoretical help and stimulus from formal discussions with Soviet colleagues encountered at international conferences.

The American side now has no access to the Soviet Union's unique deep-water facility — lake Baikal, which has a maximum depth of 1,620 metres and in many places descends to below 1km within a few hundred metres of the shore. Baikal has the further advantage that the currents and bed structure have already been studied by Soviet limnologists. In March 1982, it was announced that a special station for the recording of muons and neutrinos would

be established at Maritui, on the southwest shore of Baikal. Initially, there will be a detector lattice of up to one million m^3 , which, will "probably" be increased to 1,000 million m^3 at a later date.

Meanwhile other neutrino projects are going ahead, using, *inter alia*, a 300-metre deep salt-mine in the Donbass and a four-storey observatory inside Mount Andyrchi in the Caucasus. At the Dubna Joint Nuclear Research Institute, and the Serpukhov High Energy Physics Institute, moreover, theoreticians continue to investigate the possibility of "heavy" neutrinos, with a small, but non-zero rest mass.

Another joint Soviet-US neutrino experiment, announced in spring 1981, has failed to materialize but apparently for scientific, rather than political reasons. Code named Batiss, the experiment involved the generation of a stream of neutrinos from the Batavia research centre in the United States, and their detection beneath lake Issyk Kul in the Tien Shan mountains. In accordance with the officially utilitarian basis of Soviet science, this experiment was announced as a means of checking the hypothesis of tectonic plate movements and the early forecasting of earthquakes.

Vera Rich

Medical education

Obstacles for private school

READERS with a charitable disposition and money to spare to help buy a medical school are invited to write to Dr Paul MacLoughlin at 146 Harley Street, London W1. Dr MacLoughlin's ambitious scheme to establish a new independent medical school in central London is being held up by lack of funds, and plans to open the school in October have had to be abandoned.

Dr MacLoughlin had hoped to buy the building recently vacated by his old medical school, the Royal Free Hospital Medical School in Hunter Street. But fund-raising has been difficult and the building remains unsold. The latest of several feasibility studies, all paid for by Dr MacLoughlin, will be completed shortly and may enable him to secure a bank loan to cover the purchase price.

Dr MacLoughlin is at pains to emphasize that the projected school, to be called the Hunter School of Medicine, would not be a profit-making venture. The school has the status of a registered charity and this is proving to be one of the main stumbling blocks for finding sources of finance. Dr MacLoughlin feels, however, that only an independent school will allow him to develop his unconventional ideas on training and student selection. He says that many potentially good doctors are lost to the profession because they do not have the high school-leaving examination grades in science subjects that state medical schools are able to demand. The Hunter School would take into account in its selection process other personal qualities such as motivation, willingness to take responsibility and even moral and philosophical outlook.



Desirable medical school for sale — the old Royal Free buildings. Reproduced by kind permission of British Medical Journal.

Once the building has been bought, Dr MacLoughlin thinks that further funds will be more readily forthcoming. He is already over-subscribed with potential students and has offers of money from overseas governments. His plan is to take about 40 British and 20–30 overseas students a year, each paying about £6,500 per annum.

Critics of the scheme draw attention to the high cost of providing clinical experience in such medical specialties as neurosurgery and cardiathoracic surgery. But Dr MacLoughlin claims to have many offers from doctors in National Health Service (NHS) hospitals who would be willing to provide teaching facilities. Students would be rotated around different hospitals in the area, returning to Hunter Street for verbal instruction. Dr MacLoughlin is also confident that physicians working outside the NHS would make teaching facilities available, and is not concerned by fears expressed in some quarters that private patients would be unwilling to be guinea pigs for medical students.

Money apart, one major hurdle will be recognition of the school's qualification by the Education Committee of the General Medical Council (GMC), which has statutory powers to monitor and advise on approving medical teaching and examinations. A further snag is the European Community directive that medical education must be supervised by a university. The would-be Hunter School is looking towards the privately-run University of Buckingham, although no formal approach has yet been made.

The university itself has been considering offering pre-medical courses and may later consider collaboration with the Hunter School on pre-clinical teaching. As GMC would have to monitor an entire generation of students before it could approve the Hunter School's course (and the council has no precedent for validating an independent school) approaches have now been made to the ancient Society of Apothecaries, which still has the power to approve names for the medical register.

Despite Dr MacLoughlin's enthusiasm, several independent authorities have doubts about the scheme's financial viability and the need for more qualified physicians. The Association of University Teachers has condemned the plans on the grounds that while school-leaving qualifications "have their limitations, the cheque book is not a suitable alternative".

Despite general pessimism about the project's chances of success, the association says it is not taking any chances in the current political climate. Dr Bill Stephenson, of the association's London branch, pointed out that 12 years ago people had thought the idea of an independent university inconceivable. But the University of Buckingham is now chartered to award its own degrees and has the Lord Chancellor, Lord Hailsham, as its first chancellor.

Tim Beardsley

French patents

Fanning fires of invention

THE French minister of industry and research, M. Laurent Fabius, believes that French genius is not protecting itself sufficiently with patents. So earlier this month, he proposed a 20-point programme to the council of ministers designed to encourage French men and women to embarrass the French patent office — the Institut National de la Propriété Industrielle (INPI) with inventions.

As things are the French balance of payments in royalties is markedly negative. France pays out FF3,500 million (£286 million) each year in royalties on foreign inventions licensed in France. But it receives only FF2,100 million (£171 million) in royalties on French inventions licensed abroad. The net loss of £115 million a year indicates, in Fabius's view, both excessive technological dependence on foreign invention and insufficient French protection.

Analysis of numbers of patents (see figure) speaks equally loudly. In 1980 France granted 10,000 patents to its citizens, and 35,000 — more than three times as many — to foreigners. In Japan, by contrast, 160,000 patents were granted in 1980 to Japanese inventors but only 30,000 to foreign inventors.

The patenting process in these two countries is roughly comparable. There is essentially no assessment or weeding out of applications in either country — anyone who applies (and pays) receives. By contrast West Germany's system is much more rigorous, so that the 50,000 patents to German nationals and 18,000 to foreigners probably represent a real national advantage. Britain's system falls in rigour between the French and the German and produced 19,000 patents for Britons and

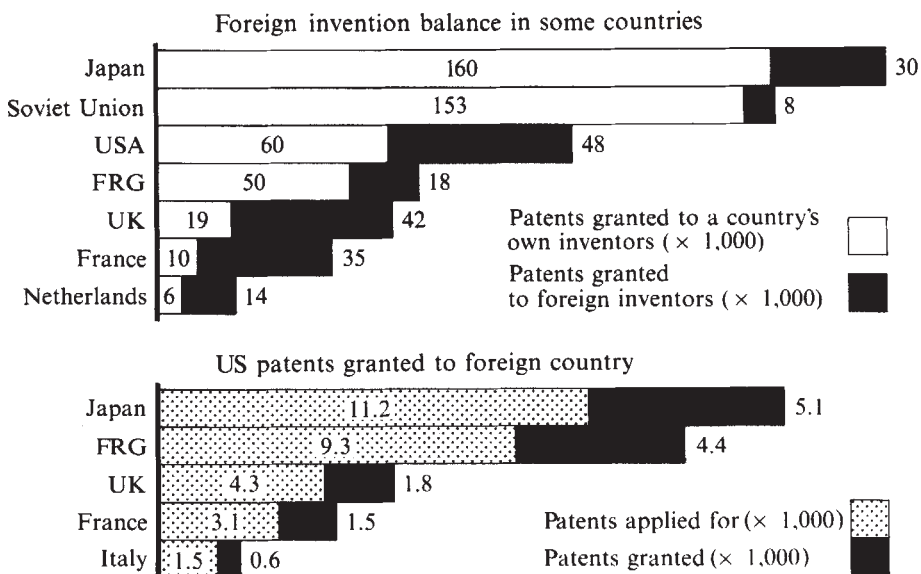
42,000 for foreigners in 1980.

A more objective test involves comparison of the number of US patents granted to certain foreign countries in 1979 (see figure). A study by Sally Wyatt and Luke Soete of the Science Policy Research Unit (SPRU) at the University of Sussex appears to show the underlying strength of West Germany in patenting — it follows close behind Japan on these figures — and the relative weakness of other countries in Europe. However, the US patents system is somewhat like the British — an application is examined and judged only if challenged. Thus the apparent patenting strength of a country is the product of a number of factors, high among which is the sheer volume of applications. "The Japanese will patent absolutely anything", Sally Wyatt claims.

But however difficult it is to interpret the data, the SPRU authors argue that they are useful among a clutch of indicators — including, say, research expenditure — which must be taken together to indicate the technological strength of a country. And M. Fabius seems determined to get this particular indicator up.

His 20 proposed measures include: support to small and medium industries to reduce the cost of patenting, now estimated to be FF 150,000 (£12,000) for world rights; free legal advice; research organizations to pay increased grants to groups which deposit patents; new courses on patenting at leading engineering schools (the *grandes écoles*); and a publicity campaign to encourage patenting, to be launched in the autumn. However, he has not been able to propose a reform of French patent law, which is little changed since the revolution.

Robert Walgate



Sources: Top, World Intellectual Property Organization. Bottom, S. Wyatt & L. Soete, *Scientometrics*, Vol. 5, 31–54 (1983).

UK research economies

British weather to cost more?

THE Meteorological Office is the latest British government agency to come under the scrutiny of civil servants briefed to find ways of increasing efficiency. But it appears to have survived the ordeal better than some, and the implications for research staff are less than catastrophic, although its customers seem likely to have to pay more in future for services provided. The Met. Office's response is not expected for some months.

The *Resource Control Review of the Meteorological Office*, produced by the Management Services Division of the Ministry of Defence, has some complimentary things to say about the Met. Office's international scientific standing, which it is keen not to damage. The Met. Office plays an important part in the World Meteorological Organization and has recently been designated one of the two World Area Forecasting Centres (the other is in Washington, DC). As well as providing weather forecasts for military and civil users, the Met. Office spends about £8 million a year on research in meteorology and geophysics, with a high proportion geared to customer needs.

The review does, however, conclude that there is not enough external scrutiny of research activities, and urges reconstitution of the Meteorological Research Committee, composed of Met. Office staff and independent experts, which was abolished a year ago as an economy measure. Indeed, the review says, the committee should have increased powers to advise on research projects and to establish collaborative links with universities, research councils and other external organizations. The review also suggests restoring the research fellowships for visiting academics which fell by the wayside when manpower counting rules were altered so that fellowships were included within staff ceilings.

One of the review's most significant recommendations is that the director general should have greater delegated financial powers from the Ministry of Defence, with lower levels of management taking more responsibility for planning. Bureaucratic limitations are felt by some to have hampered the Met. Office in the past. Less popular will be the suggestion that the training college, Shinfield Park, might be replaced or at least part of the grounds sold off. Staff levels have fallen by 13 per cent in the past four years and the college is now under-used.

Most of the potential savings identified concern the provision, and, particularly, the charging policy for forecasting services. It is recognized that the state meteorological service should continue to provide some free forecasts, to the press and to the British Broadcasting Corporation, but the review says that the ultimate objective for

all other services should be to recover full costs. At present the philosophy behind the Met. Office's charging policy is to maximize revenue, and it is feared that dramatic price increases for specialized services, such as those provided to farmers and builders, might deter users for whom forecasts are not absolutely essential.

One specific recommendation the review makes is that it should no longer be possible for members of the public to have free telephone access to a forecaster. It also calls for a review of the level of forecasting provided for Royal Air Force stations in Britain and suggests that savings could be made by grouping stations together for this purpose. The Royal Air Force, it can safely be assumed, will resist any such attempts.

Sir John Mason, director general at the Met. Office, says that many of the recommendations will have to wait until a formula for cost accounting to the different services is agreed. At present, costing is done on the basis of staff costs but with increasing automation this formula has introduced distortions. The review ducks this problem, suggesting that it be tackled by another committee.

Tim Beardsley

BA evolves

THIS year's annual meeting of the British Association for the Advancement of Science, to be held next week at the University of Sussex, marks something of a departure from the format of previous years. Sir John Mason, the association's president, hopes that the changes will attract more practising research scientists to the meeting and aims to turn it into a "parliament and festival of British science and technology".

The major innovation is a series of specialist seminars, which are expected to attract several hundred research scientists. The seminars, called "Mason conferences", have been organized by various learned societies, which have been given a free rein in arranging topics. This year most of these conferences are in the physical sciences but if the idea proves a success the list will be extended next year to include biological and social sciences.

Biologists will not be short-changed, however — biotechnology is the subject of a day-long general symposium. Many of the talks in other subjects will be linked to the four themes for this year's meeting — disasters, science in Europe, science policy studies, and land use and resource exploitation. Children are catered for by a special programme of activities and there are also events for young people. The organizers expect, on the basis of advance bookings, that the meeting will be crowded.

Tim Beardsley

Soviet industry

The old old problem

A NEW "experimental" scheme of financial incentives announced last month marks a renewed assault on the besetting problem of Soviet industry — how to implement research results rapidly at shop-floor level. During the past 15 years, successive five-year plans have called upon scientists to provide the production sector with technological innovation, and have urged the establishment of new structures such as regional science centres to encourage closer cooperation between research institute and shop-floor or collective farm. These efforts, however, have been in many cases vitiated by factory managers, who, under the present system of calculating wages and bonuses, see the new technology as a threat to their income.

Soviet planning, with its rigid system of quarterly, yearly and five-yearly targets, makes no provision for temporary shut-downs for modernization. While it is relatively easy, therefore, to put new technology into a new plant, the renovation or modernization of an existing plant imposes considerable financial penalties on the work force. Not only is there a temporary loss of bonuses, due to failure to achieve targets that in many cases were set before the need for modernization became apparent, but, in the case of energy and resource-saving technologies, the wage-fund may even be permanently reduced.

Although there are some financial incentives for innovation and resource saving (including an elaborate system of medals and orders), hitherto the wary manager would make his own private cost-benefit study before deciding whether to support or frustrate efforts to install the new technologies proposed by well-meaning scientists. Sometimes, it appears, plant managers have even entered into contracts with research institutes to undertake some project, as urged by the directives on "scientific and technical progress", with no real intention of ever acting on the results.

The new procedures, which at present are confined to factories under the ministries of heavy industry and the electro-technical industry, together with selected light-industry plants in the Byelorussian, Ukrainian and Latvian republics, represent the first real effort to come to grips with this problem. Accordingly, in addition to changes in the system of indices used for calculating wages and bonuses (so as to act in favour of innovation), managers will be given a considerable degree of independence in the allocation of funds earmarked for research and development. These can now be used not merely for research, design and pilot studies, but also to cushion the short-term disadvantages of introducing new technology.

Vera Rich

Too much high-energy physics

SIR — Your Washington correspondent reports (*Nature* 9 June, p.465) that Dr Keyworth, science adviser to President Reagan, upon being asked about the very strong adverse reaction of the entire materials research community to the Lawrence Berkeley Laboratory (LBL) proposal for a National Center for Advanced Materials (NCAM), agreed that it was not the optimal step for the health of materials science, "but the issue is what is best for America in toto? Where you allocate federal resources is not a scientific issue but a matter of priorities".

I have had the rare privilege of spending a year as a scientist thinking precisely about the allocation of federal resources in the technical sector within an organization studying national policy making without the constraint of day-to-day decision making.

From that viewpoint this administration's science-policy decisions are a baffling crazy-quilt. My perception of the national priority is to reconfigure the total research and development effort into a much more efficient system to improve the rapidly declining position of the United States in useful science of all kinds and in most technologies, with the single exception of defence technology where the United States remains far ahead in sophistication. Restoring the health of the "midday" industries (those which are essential and used by large numbers — food, materials, housing, transportation) which provide the main industrial employment base while also encouraging the "sunrise" ones could certainly constitute the focus of technology policy. Basic non-purpose-linked science and our highest tech industries (Bell Labs, IBM and so on) are the bright spots in the American scene — probably needing the least help.

Yet the policies have been almost exactly the opposite. Destabilizing of the fragile bureaucracy for applied science by successive proposed and planned abolitions of the Departments of Energy, Education and Commerce. Proposed abolition of almost the best models of essential government-led laboratories helping industrial research — the Fire and Building Centers at the National Bureau of Standards. Abolition of the National Science Foundation's science education directorate and at the same time enormous (~20 per cent) proposed increases in basic science budget and special initiatives such as LBL-NCAM.

Analysis of this pattern of "allocation of federal resources" clearly says the following about the Reagan Administration's priorities. It has no responsibility for the saving of American technology and the jobs dependent on it, it has no responsibility for either creating or partly redeploying science technical personnel away from the more esoteric sciences (presently forced by the federal budget allocations) to the equal-

ly basic, considerably more difficult science relevant to real private and public sector technologies. The reasons for this, I believe, are human. For far too long the national science priorities in the United States have been made by the high-energy physics community. By sheer numbers, seniority, gratitude for the bob, false hopes in the star war victories, to say nothing of hubris, this community has rephrased General Motor's Engine Charlie's aphorism to read "what's good for physics is good for science and the nation".

Perhaps LBL-NCAM can become the first crack in this paradigm by proving that what may be good for physics has very little to offer materials science, and right now materials research is vastly more connected with national priorities of a healthy technology and more jobs than a sixth synchrotron light source. One lesson must not be forgotten by the materials community and all other applied science and engineering groups: if they don't aggressively make policy in their own fields for the national good, others innocent of any experience or knowledge of the field will do it for them.

RUSTUM ROY

*The Brookings Institution,
Washington DC 20036, USA*

Farmers' help

SIR — *Nature* is to be commended for its important editorial (23 June, p.645) concerning agricultural research policy in the United States. By pointing out that genetic science has the potential for reducing costs as well as increasing crop quantity, it hits a neglected nail squarely on the head and highlights new opportunities for improving farmers' income. This can help move urban-based policy makers beyond a simple minded focus on "yield".

Consider a plausible illustration: semio-molecules (loosely, "signalling-molecules") are produced by many plants and operate in a pheromone-like manner by causing insects and other pests to avoid the plants for feeding or egg laying. Suppose that functionally active semio-molecules, which are also strongly promoted single-gene products of suitably low molecular weight, can be ubiquitously expressed in host plant tissues through genetic engineering. Such an outcome in pest management via plant molecular and cellular biology could be several orders of magnitude more cost effective than traditional chemical approaches to formulating, manufacturing and applying biocides. A genetically engineered "non-preference" might also present a safer approach to the complex public health issue of toxicity.

While a number of challenges for genetics and ecology are raised by this hypothetical illustration, several interesting points could be made: (1) this type

of technology is now foreseeable; (2) some polygenic hurdles might be side-stepped; and (3) the added value in this case may be one of cost reduction (eliminating the need for insecticide application), not of increased harvest index.

Other examples with far greater potential for farmers can be contemplated. All would require enhanced collaboration between traditional agricultural science disciplines and state-of-the-art biology and biochemistry. In any case, we should be helped if we refined our use of the word "yield" to mean Return On Investment rather than as a measure of gross biomass per acre.

R. N. DRYDEN JR
*AgriGenetics Corporation,
3375 Mitchell lane,
Boulder, Colorado 80301, USA*

Greek science

SIR — In an interview by Robert Walgate with the Minister of Research and Technology, Professor G. Lianis, Mr Lianis states that one of his priorities will be "to sort out which scientists in the universities and the few research institutions in Greece are more or less political appointees of previous governments and which are doing real science". Thus the impression is given that the research centres of Greece are political fiefdoms.

Concerning Demokritos, the largest of the "few research institutions", the statement of Minister Lianis is an accusation to which we vehemently object. The scientists employed at the Demokritos Nuclear Research Centre have been chosen on purely scientific criteria. Additionally, for at least the past fifteen years, all scientists at Demokritos have been appointed after the position has been properly advertised in the press and a scientific committee has examined the qualifications of each applicant. This is a procedure required by law. Furthermore, each scientist hired for the first time signs only a one year contract that can be renewed twice more before he acquires tenure.

The innuendos that good research has never taken place in Greece create another false impression. We wish to refer you to another article in *Nature*², where a *Nature* travelling fellow visiting Greece states that "the factors contributing to Demokritos' success seem to be . . . the higher qualifications possessed by practically all scientific staff and good facilities".

We would also like to inform you that in a press release of our association published in the Athens newspaper *Kathimerini* (30 April 1983), a comparative study of Demokritos with other Western European national laboratories for the period 1979–81 reinforces the view expressed in your 1978 article. G. VOURVOPOULOS
*Demokritos Nuclear Research Centre,
Aghia Paraskevi-Attikis, Athens, Greece*

1. *Nature* 301, 363 (1983).

2. Pantelouris, E.M. *Nature* 271, 394 (1978).

Mauna Kea telescope debate

SIR — Sir Bernard Lovell (*Nature* 21 April, p.650) has given a valuable resume of the history of the UK millimetre wave telescope but he did not mention the site. From his account I learned with increased concern that the telescope is now being designed for sub-millimetre wavelength and this was confirmed by an announcement (*Nature* 16 June, p.564) that construction had begun. The telescope is to have a 15-m diameter disk accurate to the Rayleigh limit for a wavelength of 0.3mm. This is indeed a courageous enterprise and it is not surprising that the cost is to be borne by the UK Science and Engineering Research Council (SERC).

Since the telescope is now being made for sub-millimetre wavelengths, the driest site available must be used, so my suggestion¹ that a medium altitude desert site might be considered as an alternative could now be invalid. The choice of Mauna Kea could, therefore, be justified if it were proven that a telescope located there could be used with reasonable efficiency to make observations at 0.3mm. This I doubt.

The possible impediments are these. The acceptable amount of precipitable water is much lower at this wavelength than for operation at 1 mm, which may severely limit the amount of usable telescope time. In addition, an anomalous absorption component must also be absent or small. The possibility of such absorption has been shown by the only two published measurements^{2,3} of atmospheric transmission at Mauna Kea. Finally the fluctuations in transmission, which there is good reason to believe are associated with anomalous absorption, must not be such as to swamp receiver noise. These are stringent conditions and the pointers from published work suggest that they are not often realized.

I have examined most of the references given by Dr Richard Hills (*Nature* 21 April, p.650), in support of his contention that the Mauna Kea site is adequate and my interpretation of these is different from his. The conditions in which observation at the shorter wavelengths can be made at all are the exception rather than the rule. This is supported by the experience at Mauna Kea of some US workers⁴ who say:

Even from a high dry site such as Mauna Kea the atmospheric transmission at 434 μm is highly variable and acceptable for astronomical observations for only a small fraction of the time.

This work was cited by Dr Hills in support of his views. I submit that he is not justified on this basis, and indeed from other references he has quoted, in rejecting the results of ref. 3 and, by implication, those of ref. 2. These two groups of researchers were quite independent and their results are in substantial agreement.

My doubts about Mauna Kea have additional support from a recent *Nature* publi-

cation⁵ which showed an obvious gap in an interesting spectrum measured at Mauna Kea; the authors remarked:

Although the conditions were stable during the run the atmospheric water vapour was insufficiently low to allow 350 μm observations to be productive.

The question of the adequacy of the site could, however, have been completely resolved by the results of a sufficiently sustained and thorough investigation of the behaviour of atmospheric transmission for wavelengths extending down to one-third of a millimetre. If this has already been made, there can surely be no reason for withholding its results from immediate publication. If it has not, then SERC is open to the charge of irresponsibly starting to construct an expensive telescope on a site where its high precision may be effective so rarely that it will not justify its construction and maintenance cost.

The pressing need for improved justification or complete reconsideration of this project is evident from the high real cost. There is, in addition to the capital cost of £7 million, a sum to be added for maintenance throughout the telescope's effective lifetime, which could amount to about £15 million. The general issue is the justification for another £20-odd million investment in astronomy, 80 per cent of which is funded by the UK taxpayer, at a time when SERC is unable to support alpha-plus proposals in small science. The doubt is aggravated by the suspicion that in this case the project may be technically on a much less sure foundation than the various radio, infrared, optical, ultraviolet and X-ray region projects which the taxpayer has already supported handsomely. British astronomers have had a generous share of national resources. The time has come for other enterprises to have support.

H.A. GEBBIE

*Imperial College of Science and Technology,
South Kensington,
London, SW7 2BJ, UK*

1. Gebbie, H.A., Llewellyn-Jones, D.T. & Knight, R.J. *Nature* 299, 280 (1982).

2. Nolt, I. *et al. J. Atmos. Sci.* 28, 238 (1971).

3. Moffat, P.H., Bohlander, R.A., Macrae, W.R. & Gebbie, H.A. *Nature* 211, 580 (1981).

4. Fetterman, H.R. *et al. Science* 211, 580 (1981).

5. Gear, W.K. *et al. Nature* 303, 46-47 (1983).

RICHARD HILLS REPLIES:

SIR — Dr Gebbie's letter is founded on the premise that the main purpose of the UK/Netherlands telescope is to observe in the "windows" at 0.35 and 0.45 mm. This is not so. The approved specification is that it shall be "usable over the range 13 to 0.3 mm, but optimized for the range 4 to 0.8 mm". The unique combination of large collecting area, high precision and high altitude site will enable the new instrument

to undertake an enormous amount of exciting astronomy in the 4 to 0.8 mm waveband.

There is ample evidence to show that atmospheric absorption will not be a serious hindrance here. The evidence comes from published observations, of which the references given in my previous letter (*Nature* 21 April) are only a sample, from measurements made available directly to the project by both astronomers and atmospheric physicists, and from several site surveys. It is clear from these that the strong absorption reported by Dr Gebbie's group is not normally present.

The virtually unexplored waveband between 0.8 and 0.3 mm offers a tremendously challenging further opportunity. Naturally we are well aware of the difficulties caused by the atmosphere at these wavelengths, although these are somewhat mitigated by the steeply rising spectra of many of the sources to be investigated. The fact is that such observations are already being made with the existing telescopes on Mauna Kea.

The paper by Fetterman *et al.* is a fine example of what can be done: as well as spelling out the problems in the passage quoted by Dr Gebbie, they report important new findings on the high velocity gas flow in Orion. The quotation from Gear *et al.* simply confirms that the 0.35 mm "window" is not always open. Our estimate is that something like 2,000 hours a year will be suitable for observing at these shortest wavelengths, and the plans for the 15-m telescope are designed to make maximum use of these opportunities. Rapid beam switching will be used to subtract atmospheric emission and up to four receivers can be installed simultaneously so that the observing frequencies can be changed quickly in response to changing conditions.

It will be possible to make observations under remote control from sea level and even direct from Europe. This will cut costs and facilitate flexible scheduling of the telescope so that the best atmospheric conditions can be used for the most critical observations. It is in order to plan this in more detail that continuous monitoring of the atmosphere is now being carried out and, when a reasonable statistical base has been obtained, the results will be published.

Finding the right balance of expenditure between world-class research facilities such as the millimetre-wave telescope and innovative "small" science projects is an important problem, but it is quite separate from the question of atmospheric transmission on Mauna Kea. Dr Gebbie is entitled to his opinions, but it seems to me that these issues need deeper analysis than he has given here.

RICHARD HILLS

*Cavendish Laboratory,
University of Cambridge,
Cambridge CB3 0HE, UK*

The Geophysical Journal at 25

The first satellites and the Geophysical Journal were launched about a quarter of a century ago. Earth sciences continue to benefit from both.

ONE of the unsurprising oddities of British public life is that the Royal Astronomical Society should also be the publisher of one of the most widely read of British journals in the earth sciences, the *Geophysical Journal* (with of the Royal Astronomical Society added for formality). The journal is now a mere 25 years old, and has just celebrated its anniversary with a special issue full of review articles celebrating developments in the past quarter of a century (*Geophys. J. R. astr. Soc.* 74, 1-376; 1983). There is also a brief introductory note by A.H. Cook and T.F. Gaskell, the founding editors, which is both informative and laconically dissembling. The story as told is interesting; what is left unsaid is absorbing.

In a sense, there always has been a *Geophys. J.*, or there has been since 1921, but it began life as the *Geophysical Supplement* of the *Monthly Notices of the Royal Astronomical Society*. That connection is easily understood. Ever since Aristarchus (if it was he) measured the diameter of the Earth, it has been clear that there are some lines of astronomical enquiry that turn back on the Earth. With the vivid interest in the theory of the motion of the Moon, and in tidal theories, towards the end of the nineteenth century, it was natural that astronomers should have acquired an interest in the departures from spherical symmetry of celestial objects and in the distribution of mass within them, and therefore proper that they should have an interest in how things became like that.

Similarly, even after the first editor of *Nature* was shown to have been mistaken in his belief that the solar corona was a phenomenon in the outer atmosphere of the Earth, astronomers were best fitted to observe a number of terrestrial phenomena — the aurorae and phenomena observable during total solar eclipse.

But where could a mass of material such as this be published? Many fortunately chose *Nature*, others chose interdisciplinary journals such as *Proceedings of the Royal Society* and then, from 1921, there was the ponderous device of the *Geophysical Supplement*. People like the late Sydney Chapman often found it more convenient to write books instead. In Britain, however, the young Sir Harold Jeffreys seems to have been the most persistent problem for editors of publications extant half a century ago. By the standards set for themselves by card-carrying earth scientists, geologists and the like, his work on the internal constitution of the Earth was

pure speculation, devoid of observation. But plainly (as it must then have seemed) it had no place in an astronomical journal either. So in a sense there was a need of something like *Geophys. J.* to accommodate Jeffreys and his school.

By 1958, there were two other developments to contend with — the growing belief that continental drift might be real after all, based largely on palaeomagnetism, which was unpalatably radical for the geological journals, and the rich harvest of the International Geophysical Year.

So was it mere entrepreneurship that persuaded an astronomical society to start a geophysical journal? There is more to the tale than that. A quarter of a century ago, the standard British journals in the earth sciences were unsympathetic and were profoundly suspicious of the people who held that it should be possible to find out about the structure of the Earth with the help of instruments other than the hammer and the microscope. *Geophys. J.* is thus a monument, probably permanent, to an old quarrel which has largely been settled but which is far from forgotten.

Cook (still managing editor of the journal) and Gaskell admit that at the outset, luck was on their side. The launching of the first two sputniks helped them to sell subscriptions to the new journal and also gave them an unexpected stream of material to publish. Fifteen years went by before the geologists were reconciled to the theory of plate tectonics, and one of the persistent ironies is that Jeffreys, whose activities helped to make the need for the journal apparent and whose 70th and 90th birthdays have been celebrated by special issues of *Geophys. J.*, remains disaffected.

The objective of the latest celebration is different — to show how geophysics has changed in a quarter of a century. Cook and Gaskell insist on the importance of improved techniques, and they are surely right. The new techniques include those for observing the motion of Earth satellites, to begin with by straightforward (but far from simple) analysis of their orbital motion, more recently by laser tracking and by radio altimetry from the satellites themselves. King-Hele, from the Royal Aircraft Establishment at Farnborough in Britain was one of the first to recognize what might be wrung out of other people's satellites by careful observation, for which reason it is proper that the first article in this celebratory symposium should be his.

In 1958, the geometrical shape of the sur-

face of the Earth was known to within a few metres by means of a blend of astrometry and triangulation. Now, of course, triangulation based on satellite tracks and, potentially even more important, interferometry of distant radio sources, promises to reduce the uncertainties by two orders of magnitude, to a few centimetres at most. With the help of such data the modellers of the Earth's gravitational field make use of a synthetic reference surface of the Earth called the geoid — a surface containing as much mass as the Earth itself, but which is an equipotential surface whose shape is also consistent with the pattern of the gravitational field external to itself.

The best measure of what has been learned since 1958 is provided by the maps of the shape of the geoid which K. Lambeck and R. Coleman contribute to the special issue of *Geophys. J.* (74, 25-54). They have the effrontery to draw contours at intervals of 3 metres, and conclude that the various estimates so far produced of where the geoid lies are discordant by only a few metres. The point seems to have been reached at which the patterns on the surface of the Earth call for a more tangible interpretation, in terms of mantle convention plumes or spreading ridges, than the plate tectonics people have provided.

The chief interest of this field, however, is still largely technical — how to extract from the data which have accumulated a more complete description of the Earth's gravitational field? Lambeck and Coleman demonstrate convincingly that the tracking of a diversity of satellites, with different altitudes and different orbital inclinations, is essential. It is remarkable how much knowledge of high-order terms in the spherical harmonic of the gravitational field King-Hele and others have extracted from the chance coincidence of some satellite orbit with some resonance caused by the Earth's rotation, bringing some gravitational anomaly around so as to affect a satellite orbit cumulatively.

Yet Lambeck and Coleman's account suggests that the analysis of the data now available is not nearly as systematic as it should be. People seem repeatedly to analyse data provided by a new clutch of satellites, to publish a new Earth model and then turn their attention to something else. Is it too much to hope that somebody will devise and maintain the computer program that will enable this important job to be done systematically?

John Maddox

Cancer

Step by step into carcinogenesis

from John Cairns and Jonathan Logan

MOST human cancers develop as the result of a succession of steps occurring over a period of many years. For example, the first visible step in the production of a cancer of the cervix is commonly the appearance of an expanding clone of abnormal cells (cervical dysplasia); the clone usually regresses and disappears, but sometimes it gives rise to a still more abnormal family of cells (carcinoma *in situ*), and it is among these that the fully invasive cancer arises. The sequence therefore involves the release of cells from the physical restraints that operate in an organized epithelium, a relaxation of the rules of differentiation and exemption from the rule of finite lifespan. Since we do not understand the molecular biology of the growth and form of even the simplest multicellular system, it is impossible to guess what changes in biochemistry would endow a family of cells with that rather abstract list of properties. In particular, we have no way of guessing whether cancers arise in many different ways, sometimes involving one set of gene products and sometimes another, or usually follow a well defined course involving the same set of functions.

It was therefore a great simplification to discover, during the last year, that some human cancers are abnormal in the same functions that are affected by the RNA tumour viruses of animals. In one kind of human lymphoma and in certain leukaemias, which were known to be associated with particular chromosomal rearrangements, the change was shown to involve the cellular homologues of certain known retrovirus oncogenes (proto-oncogenes). At about the same time, DNA from several human epithelial and lymphoid cancers was shown to be capable of transforming mouse cells in culture (though the majority of cancers which have been tested so far do not yield DNA active in this assay). In some instances this transforming DNA has also been identified as one of the known proto-oncogenes, behaving abnormally because it is overexpressed and/or bears some alteration in coding sequence. So there was the pleasing prospect that apparently diverse modes of carcinogenesis may actually proceed through a limited set of cellular pathways. Three papers in this issue of *Nature*, on pages 596, 602 and 648, extend the process of simplification at least one step further and also cast some light on the classes of function that have to be altered to make a cancer cell.

Originally, it was possible to identify oncogenes involved in certain cancers and to determine the localized sequence changes responsible for their activity, because the altered DNA, behaving like a dominant

trait, could be incorporated into other cells by transfection and would make them tumorigenic. It was customary to use NIH 3T3 cells for the experiments because their properties had been well studied and they had proved to be very efficient recipients of transfecting DNA. This is, however, a line of mouse fibroblasts that have been 'immortalized' by passage through crisis in culture; furthermore, these cells occasionally undergo spontaneous transformation, suggesting that they may have already passed through all but the final step of carcinogenesis. So they may provide a very restricted assay for the genetic changes present in cancer cells. One of the built-in programmes protecting the whole organism against invasion by clones of uncontrolled variants is the programmed senescence that normally comes into play after a certain number of cell generations. This barrier has to be overcome whenever a cancer is formed or a cell line is established *in vitro*, and so it is not logical to expect that transfection experiments with any established cell line could show what changes are needed to circumvent programmed senescence. The key ingredient in the new experiments is the use of primary and secondary cultures of rat and hamster cells to observe what phenotypic changes are produced by tumorigenic oncogenes and to see whether it is possible to transform these more normal cells using some combination of oncogenes.

A year ago we learned that transfection of a gene isolated from the human bladder carcinoma cell line EJ will transform NIH 3T3 cells. This gene (Ha-ras), one of several related *ras* genes dispersed among the human chromosomes, is a homologue of the Harvey murine sarcoma virus oncogene and this version of the gene was tumorigenic for NIH 3T3 cells because it bore an alteration in one codon. The present set of experiments establishes that this mutant gene will not, by itself, transform

normal rodent fibroblasts (though it can induce the ability to grow in soft agar). However, it will transform if the recipient cell is, at the same time, given a virally promoted *myc* gene (see Land, Parada and Weinberg, p.596) or certain genes from DNA tumour viruses (Land *et al.*; and Ruley, p.602), or if the cell has already been treated with carcinogens and has undergone immortalization (Newbold and Overell, p.648). For tumorigenicity, it is therefore necessary to alter at least two separate functions. There were, however, signs that sometimes a third step is necessary, because fibroblasts transformed with mutant *ras* together with virally promoted *myc* produced tumours (sarcomas) with a limited capacity for growth, whereas transfection of an established line of rat fibroblasts produced tumours that would grow without limit. In short, we seem to have here a sequence of two or three steps that would nicely fit the fact that the incidence of sarcomas in human populations increases roughly as the second power of age.

At this point, a sceptic could argue that even though the number of steps agrees with the epidemiological facts, the agreement might be misleading. The easiest way of making a sarcoma by transfection may well be to modify the functions encoded in *ras* and *myc*, but the usual way sarcomas arise 'spontaneously' could conceivably be by some other route. It is therefore very significant that the role of altered *ras* can be filled by the gene of polyoma virus that codes for middle-T antigen, and the role of altered *myc* can be filled by the polyoma gene for large-T antigen or by the *E1a* gene of adenovirus 5 (pages 596 and 602). As its name implies, polyoma can cause tumours in many different kinds of cell. So it is important news that a DNA virus, generically unrelated to the retroviruses in which the *ras* and *myc* oncogenes were first discovered, apparently produces cancers by achieving the same objectives, perhaps by modifying the very same two functions that had to be modified in the transfection experiments that use cellular oncogenes. It seems quite likely therefore that two critical steps in the formation of many human cancers — even those arising in completely

100 years ago

From *Nature* 28, 397, 23 August 1883.



Auroral Observatory at Kautokeino.

different tissues — may prove to be changes affecting what we might loosely call the *ras* and *myc* functions. Which of these steps usually occurs first will depend on which, if either, confers some survival advantage to cells within a tissue.

The evidence so far is that activation of cellular *ras* proto-oncogenes in naturally occurring cancers can be achieved by altering a single amino acid — which can result from changing a single base pair — and that *myc* activation involves a DNA rearrangement event. This could mean that at least one mutational event and one transposition or rearrangement event must occur in the natural history of many cancers. We have so far no direct evidence showing how to relate these events to phenomena such as initiation and promotion; nor can we say whether these steps are among those driven by environmental factors. But the way is open to determine whether mutagens in our food, for example, drive accumulation of the kind of mutations needed to activate *ras* genes; or whether oxidizing agents and protease in-

hibitors influence the chromosome rearrangements implicated so far in overexpression of the *myc* gene (in these experiments, and in the creation of certain human lymphomas).

It is not clear how many of the controls of cell behaviour have to be altered in order to create the commonest forms of cancer. Since most epithelial cells probably have a more complex programme of cell renewal than fibroblasts, and most epithelial cancers increase in frequency with age more steeply than sarcomas, more than two or three steps may be involved in the progression of a normal epithelial cell to a cancer. Eventually, transfection experiments will be done using epithelial cells, and we shall then see whether the additional steps involve still other members of the family of oncogenes that were originally revealed by the retroviruses. □

John Cairns and Jonothan Logan are in the Department of Cancer Biology, Harvard School of Public Health, Boston, Massachusetts 02115.

Geology

Swedish chronology revisited

from D.J. Schove and R.W. Fairbridge

In the 1940s the Swedish Time Scale — based on a varve count over 14,000 years — was the standard chronology for all parts of the world. From 1950 to 1980 its place was taken by radiocarbon. However, solar fluctuations¹ and associated Holocene wiggles² soon emphasized the need for a distinction between ¹⁴C dates (b.p./b.c.) and tree-ring dates (BP/BC). The Bristlecone pine chronology³ has enabled us to calculate the differences back almost to 9,000 yr BP or 7,000 BC, but before that we still need a scale reliably based on annual counts of varve layers. Some participants at a recent conference* suggested that in Sweden a reliable revision of the Swedish Time Scale will soon provide us with an accurate chronology of some 14,000 years.

There are several difficulties with the old Swedish Time Scale. First, the latest historical varves to be counted⁴ were given an estimated medieval date now known to be too late. Second, about 10,000 Holocene varves were counted by Liden from the Late Glacial (Zone III/IV which was thus dated at about 8,000 f.Kr. in Sweden and which is indeed around 10,000 on the ¹⁴C scale). However, such varves lack the thick meltwater component provided by glaciers and varves are easily missed. Liden's measurements were neither published⁴ nor replicated. If a thousand years^{5,6} were missed, ~ 8,000 f.Kr. = 9,000 BC = 11,000 yr BP = 10,000 b.p.

Third, the Late Glacial varves measured by De Geer⁷ are not as continuous as was thought. In more continental climates such as Finland and Canada, varves are more easily counted. Cross-dating of the Swedish and the Finnish scale in 1969 led to a suspicion^{8,9} that there was a discrepancy of nearly a century in the region of the Central Swedish moraines. It was pointed out⁹ that De Geer had bridged northern and southern series by using varve series from outside Europe altogether. The search for a better 'bridge' has failed (none may be needed, D.J.S.), but other errors⁴ have meanwhile been noticed by Lundqvist^{5,6} and N-A. Mörner (University of Stockholm).

These difficulties are now being resolved. First, historical varves can now be obtained that continue to the present. A new coring method involving freezing *in situ* makes it possible to collect and measure even last year's varve: I. Cato (University of Uppsala) has thus measured and replicated series back to the medieval period. Historical series can be cross-dated with meteorological data (for example, Baltic ice data available back to 1720 discussed by L. Makkonen, University of Helsinki) to ensure that no varves have been missed.

Second, Holocene series can be cross-dated with proxy meteorological data. Back to AD 436 we thus have a summer temperature index based on Lapland tree-rings by T. Bartholin and W. Karlen¹⁰ (University of Stockholm). New varve series, with an estimated error of only 1 per

cent in the count, have already been analysed back 5,000 years and a chart of the data, exhibited by I. Renberg and U. Segerström (University of Umeå), aroused great interest. Varve thickness, organic content, mineral matter content and visual appearance provide separate parameters that can be interpreted meteorologically.

Teleconnection with the Bristlecone pine absolute scale, using, for example, the quasi-biennial oscillation, has already been achieved for Bronze Age varves in south Russia and for tree-rings in Turkey¹¹. Similar methods will make it possible to correct whatever error exists in the dating.

Finally, glacial varves from the early Holocene have also been measured and the measurement replicated carefully¹² using a new zero at Doeviken in northern Sweden. This Swedish zero can be equated^{13,14} with varve 1166 or 1167 in the reliable Canadian (Timiskaming) varve series. Cross-dating the subsequent Canadian varves with further Swedish lake varves being investigated by Renberg will make exact dating possible through the early Boreal period (Zone V).

In the south of Sweden where new floating chronologies are being obtained, thick annual (proximal) varves can be subdivided into diurnal varves recording the number of days in the pre-Bölling summers. Such a series should correlate with the floating chronology of the varves in southern New England.

In central Sweden the failure to find a bridge is no longer a problem if the Finnish chronology^{9,13,14} is accepted: none is needed. Other studies, including palaeomagnetism, dating of shorelines and oxygen-18, are being carried out by Mörner and others. Many papers provided precise radiocarbon chronologies for other parts of the world, including the Tropics and Canada. For instance, logs from Siberia were transported north of Lapland to dated Spitzbergen shorelines not in warm periods as previously thought but in cold periods; as A. Häggblom (University of Stockholm) discovered they were transported not by water but by ice. □

D. J. Schove is at St Davids College, West Wickham, Kent BR4 0QS and R. W. Fairbridge is in the Department of Geological Sciences, Columbia University, New York 10027.

1. Schove, D.J. *Sunspot Cycles*, 337 (Hutchinson Ross, Stroudsburg, 1983).
2. Fairbridge, R.W. *Nature* **292**, 670 (1981).
3. Farguson, C.W. & H.E. Suess in *Radiocarbon Variations and Absolute Chronology* (ed. Olsson, I.U.) 237 (Wiley, London, 1970).
4. Fromm, E. in *Radiocarbon Variations and Absolute Chronology* (ed. Olsson, I.U.) 163 (Wiley, London, 1970).
5. Schove, D.J. *Sunspot Cycles*, 337, 361 (Hutchinson Ross, Stroudsburg, 1983).
6. Lundqvist, J. *Boreas* **4**, 47 (1975).
7. De Geer, G. *Geochronologia Suecica Principes* (Stockholm, 1940).
8. Schove, D.J. *Geogr. J.* **135**, 598 (1969).
9. Schove, D.J. *Geogr. Annal.* **53**, 227 (1971).
10. Bartholin, T.S. & Karlen, W. *Dendrokronologi i Lapland, Dendrokronologiska Sällskapet*, Stockholm **5**, 1 (in Swedish).
11. Schove, D.J. *Sunspot Cycles*, 351 (Hutchinson Ross, Stroudsburg, 1983).
12. Bergström, R. *Sver. Geol. Unders. C634*, 77 (1968).
13. Schove, D.J. *Geogr. Annal.* **53**, 230 (1971).
14. Schove, D.J. & Fairbridge, R.W. *Ice-cores, Varves and Tree-Rings* (Balkema, Rotterdam, 1984).

*'Climatic change and related problems' was the theme of the Second Nordic Symposium held at Stockholm in May. The proceedings will be published in 1984.

Geosciences

Bumps and bolides in the history of the Earth

from Philip Campbell

HOW smooth or how spasmodic is the history of the Earth? This elusive question confronted fifty or so scientists, drawn from a variety of disciplines, at a recent workshop in Berlin*. By the end of the workshop, the participants had decided that its original title, as expressed in the first sentence above, begged too many questions to be answerable, but the original was more suggestive of the perspective provided by the workshop: that of a system in which slowly evolving processes in the interior of Earth may, in the right conditions, produce the occasional rapid variations in the oceans, atmosphere and biota that are inferred from the geological record. Considerable discussion was also given to the possible role of impacting bolides (comets or asteroids) which have recently achieved respectability as a possible explanation for rapid change.

Unsurprisingly, the paradigm underlying all the discussions was plate tectonics, wherein oceanic crust is formed, following the rifting of continental crust, by outpouring of basalt lavas from mid-ocean ridges which subsequently spread and cool to form the sea floor. Over a time scale of 10^8 years the ocean floor is then subducted beneath a neighbouring plate margin. During this process, less dense magmas differentiate from the subducting material and rise to form volcanic regions at the surface above, either on continents or as erupted island arcs ('active margins'). This newly formed continental crust, being less dense than the underlying mantle, remains at the surface, to be subjected to erosion and possibly to lateral pressures from plate collisions. The latter may lead to the thickening and folding of the crust by which mountains are formed (orogeny).

The only available source of energy for such processes is the heat that escapes from the Earth and which is produced by the cooling of the interior since the initial accretion 4.6 billion (4.6×10^9) years ago and by radioactive decay of (principally) uranium, thorium and potassium. F. Richter (University of Chicago) outlined the processes by which the Earth's interior might accommodate this flow, making it clear that, at all times, a convecting mantle underlying the crust must have existed.

The depth to which mantle convection extends remains open to debate. An important and uncertain parameter is the temperature-dependent viscosity profile of the mantle. Experimental constraints include recent rare gas isotope data, which support those thermal models that invoke two sep-

arate horizontal layers of convection within the Earth (R. K. O'Nions, University of Cambridge); the separation depth remains uncertain however.

As the aim of the meeting was to examine the various *tempi* of Earth history, Richter emphasized critical time scales of mantle convection to be borne in mind. The shortest of these was the convective turnover time, estimated as 10^8 years or so under an ocean plate. This value depends on a number of assumptions — particularly the viscosity of the zone immediately beneath the plate — and represents a lower limit. During the Archaean (>2.5 billion years ago) the terrestrial heat flux was greater and the turnover time could have been an order of magnitude smaller.

Within the framework of plate tectonics therefore, major crustal processes are driven by a continuous gradually decreasing flow of heat from the Earth's interior whose effects on the surface features are modulated by mantle convection on various time scales. In trying to place rapid change (or even catastrophes) in such a framework, the changing behaviour of continents and oceans needs to be considered. One fundamental area addressed at the meeting, therefore, concerned the production and evolution of continental crust.

Isotopic dating and geochemistry indicate that the production of continental crust has occurred episodically over the aeons. S. Moorbath (University of Oxford) and A. Knoll (Harvard University) discussed the isotopic and sedimentary records respectively; both records strongly suggest that major crustal growth took place on Earth between 3 and 2.5 billion years ago. The formation time of the Precambrian continental terrains (those that survive, at least) appears to have been 100 million years at most. Moorbath emphasized that primary ages could only be reliably obtained from Rb-Sr, Sm-Nd and Pb-Pb whole-rock isochrons, and that the global coverage using such techniques is so far very patchy. Hence the widely quoted periods of major 'pulses' of crustal growth (2.8–2.5, 1.9–1.6, 1.2–0.9, 0.5–0 billion years ago), which are indicated in age histograms using such techniques, are by no means conclusively defined — indeed there is growing evidence that, although episodic, crustal growth periods were not globally synchronous.

These are processes that seem to take place over the medium to long term (in geological terms). Before discussing the much faster rates of change that can be seen in the geological record, it is important to ques-

tion whether our plate tectonic framework is applicable at earlier times in the Earth's history. One group of participants adopted a uniformitarian approach and sought conclusive evidence that crustal tectonics operated differently in the distant past from the way it does today. They concluded that the geological record provides none, although this involved some special pleading for preservation or removal. Although examples of various individual components of current plate tectonics can be detected in earlier terrains, the first period where all components are visible is at about 1 billion years ago when the first known ophiolite sequence was generated. (Ophiolites are continental features thought to be the remnants of ocean spreading centres, being characterized by similar vertical sequences of basaltic rocks.)

By the time the first shelled fossils were laid down (about 600 million years ago), the current style of tectonics seems to have been in full swing. From that time to the present (that is, throughout the Phanerozoic), the fossil record betrays at least five sudden mass extinctions of species. D. McLaren (University of Ottawa) surveyed the Cretaceous extinction together with four other major mass extinctions (Late Ordovician, Upper Devonian, Late Permian and Late Trias) and related the possible immediate causes such as ocean poisoning, changes in surface radiation, climatic change and sea-level variation to three possible ultimate causes: changes in mantle convection; variation in solar luminosity and in the Earth's orbital parameters; and bolide impacts or supernovae. (For a report of a recent meeting on theories to explain species extinction, see J.M. Diamond *Nature, News and Views* 304, 396; 1983.)

The hypothesis of a bolide impact has received a great deal of attention over the past two years since the discovery of a ubiquitous abundance anomaly of iridium at the Cretaceous-Tertiary boundary. (Iridium is depleted in the Earth's crust relative to Solar System abundances.) Contributed reviews of the geochemical evidence (K. Hsu, ETH Zurich), recent climatic modelling (O. Toon, NASA Ames) and the cratering records and orbital behaviour of the Solar System and its minor bodies (E. Shoemaker, USGS, Flagstaff) all supported the general feeling that a cometary or asteroidal impact took place at the end of the Cretaceous. However, T. Birkelund (University of Copenhagen) emphasized that in the Danish geological sections that span this boundary, complex faunal changes are seen that cannot simply be explained by a single event, extraterrestrial or otherwise. Ammonoids and belemnoids were undergoing steady decline from the early Cretaceous. Others have reported that Cretaceous calcareous nannoplankton appear to have died out in the 10,000 years following the boundary. J. Lipps (University of California, Davis) described changes in planktonic foraminifera consis-

* A Dahlem workshop on 'Patterns of change in Earth evolution' was held in Berlin on 2–6 May 1983.

tent with a change in ocean circulation and a decrease in vertical oceanic segregation at the Cretaceous-Tertiary boundary. Thus environmental stresses of different kinds appear to have been present before and after the 'event'. The feeling at the meeting was that a bolide did hit the Earth and probably delivered the *coup de grâce*, but that changes due to other causes may also have been in progress.

The bolide hypothesis is attractive and spectacular, and has stimulated much research into the behaviour of the exogenic system (sediments, ocean, atmosphere and biota) following such an event; but it is just as intriguing to wonder how an intrinsically ponderous system such as crustal tectonics might contribute to abrupt signatures in the geological record. The bolide hypothesis is a possible candidate for the four other major extinctions as well as that of the Cretaceous, but significant changes in sea level that were taking place during most of these periods again suggest that multiple causes must be sought.

One group of participants addressed themselves to exogenic instabilities related to sea-level changes and tectonics. A globally correlated change in sea level relative to land may arise from three possible causes: a change (for example during glaciation) in the total volume of liquid water on the Earth's surface, rearrangement of crust in a horizontal and/or vertical sense, and a change in the sediment volume on the sea floor (M. Steckler, Lamont Doherty). It is the relative rates of change of these factors that determine whether a rise or fall of sea level takes place.

The role of plate tectonics is significant in all three causes. The total volume of water depends on the climate, which in turn is affected by the partial pressure of CO₂ in the atmosphere and by the Earth's albedo. The former is increased by volcanism while the latter is affected by the areal distribution of land masses. Crustal rearrangement, possibly following changes in mantle convection, can involve a change in the length and cross-sectional area of mid-ocean ridges, opening or closing of straits, or raising or lowering of margins. The degree of sedimentation is related, via climate, to the amount of exposed elevated crust.

Although these processes took place over tens or hundreds of millions of years, abrupt changes in the exogenic system may have resulted. The opening of a previously isolated basin might flush salty water into the (less salty) ocean where it would accumulate on or near the sea floor. Such stagnant saline pools could hinder the recycling of those nutrients that settled within them, decreasing the ocean fertility.

Mechanisms were also discussed by which an increase in sea level might decrease ocean fertility, emphasizing that the background conditions during the transgression are crucial in assessing the effects. H. Holland (Harvard University) considered the case of an ocean with a strong shallow minimum in O₂ content.

The rate of carbon deposition on the Earth increases with photosynthetic productivity of the biosphere and with the fractions of organic carbon buried in sediments (predominantly on the sea floor). But the buried fraction of carbon will increase as the ocean becomes less oxic. A rise in relative sea level under these conditions will rapidly increase the area of shallow shelf exposed to less oxic conditions and hence will tend to increase the fraction of organic carbon that is buried. To maintain a constant rate of total carbon deposition (for geochemical equilibrium) the photosynthetic productivity will have to decrease. W. Berger (Scripps Institute) discussed alternative mechanisms by which the balance of essential nutrients such as

phosphates and nitrates might be upset as the result of a sea-level increase.

This report has highlighted only a few aspects of the workshop, emphasizing in particular the different time scales on which plate tectonics may have affected not only the terrestrial crust but also the biota that depend on it. But this selection is arbitrary, given the kaleidoscopic array of geochemical and biotic signatures in the geological record. One is left with the feeling that any set of independent observations on a geological phenomenon continues to be greatly exceeded in number by the variables required to explain it. □

Philip Campbell is Physical Sciences editor of Nature.

Precambrian geology

The most ancient rocks?

from Stephen Moorbath

In a reconnaissance ion-microprobe U-Pb age study of individual grains of zircon (ZrSiO₄) concentrates from a metamorphosed early Precambrian sedimentary rock at Mount Narryer in Western Australia, a team from the Australian National University reports in *Nature* this week¹ the discovery of four zircon grains — out of a total of about one hundred analysed — that have yielded near-concordant U-Pb ages of between 4,100 and 4,200 Myr. This would make these rocks older than any terrestrial rocks on which reliable age determinations have been made.

The four grains were all found within a single sample of the Mount Narryer quartzite from which most other zircon grains yielded ages in the range ~ 3,750–3,100 Myr, which is within the age range of the oldest known rocks *in situ* in the area. The quartzite itself was probably deposited ~ 3,600–3,350 Myr ago, and was strongly metamorphosed at about the latter time. Accordingly the authors postulate the former existence of rocks at least as old as ~ 4,100 Myr in the vicinity of Mount Narryer, which contributed some detrital material to the much younger quartzite.

There has been general agreement for some 30 years now that the Earth, and probably the Solar System, came into existence in something like their present shape and form some 4,500–4,600 Myr ago. Indirect but strong evidence comes from measured solidification ages of meteorites and some lunar rocks, and from a comparison of time-extrapolated radiogenic isotopic abundances between terrestrial and extra-terrestrial materials².

The oldest reliably dated terrestrial rocks, which have yielded concordant Rb-Sr, U-Pb and Sm-Nd ages in the range ~ 3,750–3,800 Myr (refs 3–6), are the metamorphosed supracrustal rocks of the Isua region in west Greenland. This complex assemblage of detrital and chemical

sediments, together with a variety of acid and basic volcanic rocks, was probably deposited at and around the margins of a volcanic island and in the surrounding ocean. There is no field evidence for an older continental-type basement, whilst radiogenic isotope evidence gives no hint of source materials having a significantly older crustal residence time⁷. The oldest reliably dated true continental crust — characteristically composed of 20–40-km-thick calc-alkaline granitoid gneisses — gives ages of ~ 3,700–3,500 Myr in several continental shield areas. Some recent reports claiming ages of ~ 3,800–4,000 Myr for such gneisses must still be treated, in my view, with extreme caution for technical reasons combined with interpretational difficulties, though there is no *a priori* reason for postulating the total absence of true continental crust in this age range.

This question of the antiquity of the continental crust is at the centre of much intense and fundamental debate. On the one hand, a small vociferous minority of 'steady-state' enthusiasts want to produce the total mass of continental crust at a very early stage in Earth history, possibly before ~ 4,000 Myr (ref. 8). Majority opinion, however, prefers production and growth of continental crust over geological time by continuing chemical differentiation of the mantle, although many workers agree that the maximum rate of production of continental crust was ~ 3,500–2,500 Myr ago, by the end of which time some 60–80 per cent of the present-day mass of the continental crust may already have existed^{9–11}. Furthermore, many workers (including myself) consider that thermal and mechanical conditions in the outer parts of the Earth before ~ 3,800–4,000 Myr were such as to inhibit formation of any significant quantity of continental crust by any process involving plate tectonics — which is as effective a way for producing con-

tinental crust as any we know. That does not imply that chemical and physical differentiation into core, mantle and some sort of predominantly mafic proto-oceanic crust did not occur well before ~4,000 Myr ago. Such a world-wide proto-crust, studded with countless small rifts and volcanic islands emitting voluminous ultramafic, mafic and relatively minor acid lavas, provides a very different scenario from one involving early formation of thick calc-alkaline continental crust. Moreover, chemical and petrological considerations render it most unlikely that crystallization of some very early global magma ocean could directly produce the continental crust¹².

The ANU workers¹ make no extravagant claims in this matter, and simply state that pre-3,800-Myr silica-saturated rocks must have occurred on the Earth's crust, because zircon is almost always found in such rocks. I fear that some workers may prematurely interpret this to mean 'continental' rocks. We know from Isua, however, that 3,800-Myr-old silica-saturated volcanic rocks (metarhyolites) carry rather abundant accessory zircon, whilst I have myself observed zircon in ~10-Myr-old granophyre intrusions in Iceland, which is not a continental environment by any stretch of the imagination. One swallow does not necessarily make a summer, and four zircon grains from a single quartzite sample do not necessarily make a continent.

Finally, there is still a possibility that the 4,100–4,200 Myr ages might be an artefact generated by the combined effects of uranium loss from an ~3,700-Myr-old zircon during a later metamorphic episode at around 3,300 Myr and loss of radiogenic lead comparatively recently. The authors briefly argue against this possibility, but I feel that many workers will require further convincing about this. In the meantime, I understand that the search is on for intact remnants of the 4,100–4,200-Myr-old source rocks of the super-ancient zircon grains in the Mount Narryer quartzite: that embarrassing little gap of 700 Myr between 4,500 and 3,800 Myr badly needs to be substantially narrowed down, preferably by finding some rocks *in situ*. □

Stephen Moorbath is Reader in Geology at the Department of Geology and Mineralogy, University of Oxford, Oxford OX1 3PR.

1. Froude, D.O. *et al.* *Nature* **304**, 616 (1983).
2. Faure, G. *Principles of Isotope Geology* (Wiley, New York, 1977).
3. Moorbath, S., O'Nions, R.K. & Pankhurst, R.J. *Nature* **245**, 138 (1973).
4. Michard-Vitrac, A., Lancelot, J., Allègre, C.J. & Moorbath, S. *Earth planet. Sci. Lett.* **35**, 449 (1977).
5. Moorbath, S., O'Nions, R.K. & Pankhurst, R.J. *Earth planet. Sci. Lett.* **27**, 229 (1975).
6. Hamilton, P.J. *et al.* *Nature* **272**, 41 (1978).
7. Hamilton, P.J., O'Nions, R.K., Bridgwater, D. & Nutman, A. *Earth planet. Sci. Lett.* **62**, 263 (1983).
8. Armstrong, R.L. *Phil. Trans. R. Soc. A* **301**, 443 (1981).
9. O'Nions, R.K., Evensen, N.M. & Hamilton, P.J. *J. geophys. Res.* **84**, 6091 (1979).
10. Jacobsen, S.B. & Wasserburg, G.J. *J. geophys. Res.* **84**, 7411 (1979).
11. Dewey, J.F. & Windley, B.F. *Phil. Trans. R. Soc. A* **301**, 189 (1981).
12. Hofmeister, A. *J. geophys. Res.* **88**, 4963 (1983).

Ecology

Laboratory, field and natural experiments

from Jared M. Diamond

THREE experimental approaches to understanding species abundances and distributions have for some time coexisted uneasily in ecology: laboratory experiments, field experiments and natural experiments. Recent papers by Schoener and colleagues^{1–4} constitute important advances in applying two of these methodologies, as well as in clarifying the relative strengths and weaknesses of all three.

The virtue of laboratory experiments is of course that values of independent variables are at least in theory completely controlled. The main uses of the approach have been for detailed studies of two-species interactions and of a single species' behavioural responses. (The technique is often practised on microorganisms kept in bottles, hence the ecologists' term 'bottle experiment'.) Well known bottle experiments include those of Gause⁵, Crombie⁶ and Park⁷, who studied coexistence conditions and competition between pairs of protozoan or invertebrate species utilizing similar foods. Sheppard⁸ and Heller⁹ used laboratory cages to study aggressive interactions between two chipmunk species which field observations had suggested to be competitors. Donald¹⁰ grew two plant species in pots with aerial or soil partitions in order to separate the contributions of root and shoot interactions to competitive depression of growth.

A major weakness of this approach in ecology is that most ecosystems depend on so many species and physical variables that laboratory models become uselessly unrealistic. For example, the bottle experiments of Gause and successors told nothing about whether competition is important in natural populations. It is not even possible to know what variables are worth putting into one's bottles until their importance in nature has been established by other means. Furthermore, for most species it is difficult or impossible to keep populations alive and reproducing for many generations under laboratory conditions ('Competition between condors and vultures: a 10-generation cage experiment').

Field experiments are more realistic but at the cost of experimental control. In a typical field experiment a single variable is manipulated — a particular species is removed or fenced out, or a new one is locally introduced (for example, to a small island). The experimenter controls nothing else after selecting manipulated and unmanipulated sites.

One of the first field experiments was by Darwin¹¹, who demonstrated that mowing

or introduction of grazing animals increased plant species diversity on a lawn (by preventing some species from outcompeting others). The most extensive introduction experiment is by Schoener and Schoener³, who placed lizards on 30 Bahamian islets without lizards and thereby demonstrated that lizards could reproduce successfully on islets much smaller than those supporting lizards naturally.

The weaknesses of field experiments are at least fivefold. First, the outcome of the experiment may vary with year, season and location, because the outcome depends on uncontrollable variables such as weather and uncontrolled predators or competitors. For example, removal of desert lizards^{12,13} and rodents¹⁴ either succeeded or failed completely to affect abundances of competing species, depending apparently on rainfall and hence food availability in the particular year of the experiment.

Second, most field experiments are not run for enough generations of the species studied to test for the possibility of such variation (see discussion in ref. 1). Often, the outcome cannot be cured by diligently running the experiment a few more years before publishing. For example, the reason why natural lizard populations are lacking on small Bahamian islands where experimental introductions of lizards routinely succeed is that hurricanes wipe out the occasional self-introduced population every century or so³. Biological systems depend on species' genetic properties, yet genetic changes following experimental manipulations may require centuries or millennia.

Third, the spatial as well as temporal scale of many important ecological phenomena bars them to experimental study. For example, distributional evidence shows that habitat fragmentation leads to selective extinction of certain species^{15–19}. This process can be demonstrated experimentally in very small areas within short times (for example, by counting insects on pruned mangrove trees for several years²⁰). However, some other approach is required to understand why many bird and mammal populations failed to survive 10,000 years on large land-bridge islands fragmented off the continents by rising post-Pleistocene sea level, and to predict which bird and mammal populations will survive a long time in forest fragments made national parks^{16,17}.

Fourth, in any ecosystem consisting of three or more species the effect of sustained manipulation of species A on species B depends not only on the direct interaction between A and B but also on their interac-

tions with any other species. Thus, the experimental outcome does not permit unequivocal conclusions about the A-B interaction itself²¹. For example, if B competes with C and if A eats both B and C but prefers B, removal of predator A may paradoxically exterminate its prey C! This fact explains the result of Darwin's classic experiment showing lower plant species diversity on lawns not grazed by sheep. The same interpretation is relevant to the otherwise puzzling outcome of a more recent experiment, in which removal of zooplankton grazers dramatically increased the abundance of certain grazed algae while decreasing others²².

Finally, for many species in many places, the merits and drawbacks of field experiments become academic: local removal or introduction of species would be technically impossible, morally reprehensible and politically forbidden.

Natural experiments escape most of the drawbacks of field experiments, at the cost of sacrificing all manipulative control of variables. Instead, control is exercised solely through site selection. The experimenter's aim is to select sites that differ naturally in the presence or absence of one major factor relevant to the dependent variable but which are similar in other major factors.

Typical examples are studies comparing the abundance, morphology, and habitat range of species A on multiple islands, some of which have and others of which lack its competitor or predator species B. For example, Schoener and Toft² found that spiders were about 10 times more abundant on average on 48 Bahamian islands without lizards than on 26 with lizards, because lizards prey on and also compete with spiders. Brown²³ found that two species of chipmunk (*Eutamias*, Sciuridae) divide the forest transect altitudinally on numerous Nevadan mountains where they occur sympatrically, but that each species occupies the entire transect on a mountain where it occurs without its competitor.

Natural experiments have three types of advantage. First, they permit one to gather data far more quickly than is possible by field experiments. Thus, one can census more individuals, species, places and times, thereby increasing the scope and reliability of one's conclusions. For example, Schoener and Toft² surveyed five spider species on 93 islands with and without lizards in 20 days of field work. In the same time they would have been able to remove most (not all) individual lizards on only two islands lacking them, and they would still have had to wait up to several years for spider densities to reach new equilibrium values on the manipulated islands.

Second, natural experiments permit one to examine conditions that cannot, may not, or should not be created experimentally. For example, Brown could not have succeeded in exterminating chipmunks on a whole mountain. Furthermore, his con-

science and the US Fish and Wildlife Service would have prevented him from trying.

Finally, natural experiments reveal the end results of ecological and evolutionary processes operating over long times and large areas. On the one hand, underlying processes such as predation and competition are likely to vary in intensity among seasons or years, and it might happen that a field experiment was done at the wrong time to detect the process¹²⁻¹⁴. On the other hand, a species excluded from a habitat for many generations by a competitor or predator may lose its genetic adaptations for that habitat. Thus, removal of the competitor or predator in a short-term field experiment would not permit reoccupation of the habitat. Only the natural experiment comparing habitats occupied on islands with or lacking the competitor (or predator) for millennia reveals the niche shift²⁴.

The obvious weakness of natural experiments is that the observer does not create a known difference between two situations, but must instead decide what difference between two existing situations is the salient one. There is always the risk that some difference other than that recognized by the observer might be the salient one: some unnoticed predator, parasite, soil nutrient, or unspecified factor confined to one of the two situations^{25,26}. In science, one can never rule out the possibility that a phenomenon is due to some unspecified factor rather than to an observed correlation: the best one can do is to strengthen the observed correlation and weaken likely alternatives. For example, Schoener and Toft² used multiple analysis of covariance of spider densities on 74 islands to disentangle the effect of lizard presence from concurrent effects of island area, isolation and vegetation complexity. Yeaton and Cody²⁷ showed that song sparrow territory size on islands increased with number of competing species in a way not explained by varying food density and habitat structure, but predictable by considering the identities of the competitors and their similarity to song sparrows in ecology, morphology and behaviour. Schoener's quantitative analysis of habitat shifts in four lizard species was based on 20 sites supporting nearly all existing combinations of these species and sorted out effects of site differences in vegetation²⁸. Schoener was thereby able to determine not only by how much each species affected each other species, but also how the effect varied among age and size classes of a species. In such cases it strains one's credulity to argue that the explanation might nevertheless be some unspecified factor.

Recent studies using these three different types of ecological experiment have been useful in dispelling some earlier misunderstandings.

First, there has been concern that the choice of experimental method might be critical to the conclusions reached — especially, that conclusions about the role

of competition might be an artefact of natural experiments and might not be sustained by field experiments. In fact, of about 164 field experiments carried out by the end of 1982 to test for competition, 148 confirmed it^{1,29}. It now appears that varying conclusions about the relative roles of competition and predation in nature are not an artefact of varying experimental method. Instead, they reflect an important biological generalization about how the size and trophic status of a species, and the intensity of physical disturbance in its habitat, control its relative sensitivities to competition and predation^{1,25,30}.

Second, some proponents of each method have claimed that their method is inherently superior and is the method of choice. In fact, each method has virtues, weaknesses and limitations of scope that vary with the species studied and with the questions asked. This situation in ecology reminds one of the coexistence within atomic physics of research based on astronomical observation and Earth-bound experiments. In ecology as in other fields, availability of very different methodologies can be a source of strength rather than of disunity: conclusions tested by different methods are more robust. □

Jared M. Diamond is Professor of Physiology at the University of California Medical School, Los Angeles, California 90024.

1. Schoener, T.W. *Am. Sci.* **70**, 586 (1982).
2. Schoener, T.W. & Toft, C.A. *Science* **219**, 1353 (1983).
3. Schoener, T.W. & Schoener, A. *Nature* **302**, 332 (1983).
4. Schoener, T.W. & Schoener, A. *J. anim. Ecol.* **52**, 1 (1983).
5. Gause, G.F. *The Struggle for Existence* (Williams & Wilkins, Baltimore, 1934).
6. Crombie, A.C. *J. anim. Ecol.* **16**, 44 (1947).
7. Park, T. *Science* **138**, 1369 (1962).
8. Sheppard, D.H. *Ecology* **52**, 320 (1971).
9. Heller, D.H. *Ecology* **52**, 312 (1971).
10. Donald, C.M. *Aust. J. agric. Res.* **9**, 421 (1958).
11. Darwin, C. *The Origin of Species* (Murray, London, 1859).
12. Dunham, A.E. *Ecol. Monogr.* **50**, 309 (1980).
13. Smith, D.C. *Ecology* **62**, 679 (1981).
14. Morris, R.D. & Grant, P.R. *J. anim. Ecol.* **41**, 275 (1972).
15. Brown, J.H. *Am. Nat.* **105**, 467 (1971).
16. Diamond, J.M. *Proc. 16th int. Orn. Congr.* **616**, (1976).
17. Soule, M.E. & Wilcox, B.A. *Conservation Biology* (Sinauer, Sunderland, 1980).
18. Willis, E.O. *Proc. 17th int. Orn. Congr.* **2**, 783 (1980).
19. Whitcomb, R.F. et al. in *Forest Island Dynamics in Man-dominated Landscapes* (eds Burgess, R.L. & Sharpe, D.M.) (Springer, New York, 1981).
20. Simberloff, D. *Ecology* **57**, 629 (1976).
21. Bender, E.A., Case, T.J. & Gilpin, M.E. *Ecology* **64** (in the press).
22. Porter, K.G. *Nature* **244**, 179 (1973).
23. Brown, J.H. *Ecology* **52**, 305 (1971).
24. Abramsky, A. & Sella, C. *Ecology* **63**, 1242 (1982).
25. Connell, J.H. in *Ecology and Evolution of Communities* (eds Cody, M.L. & Diamond, J.M.) (Harvard, 1975).
26. Connell, J.H. *Oikos* **35**, 131 (1980).
27. Yeaton, R. & Cody, M.L. *Theor. Pop. Biol.* **5**, 42 (1974).
28. Schoener, T.W. *Ecol. Monogr.* **45**, 233 (1975).
29. Schoener, T.W. *Am. Nat.* **122** (in the press).
30. Hairston, N.G., Smith, F.E. & Slobodkin, L.B. *Am. Nat.* **94**, 421 (1960).

ERRATUM

In the *News and Views* article by A.L. Bloom on 'Benefits of cloning genes for clotting factors' (*Nature* **303**, 474), the penultimate paragraph should have read 'Biogenetic expression of haemostatically effective factor VIII is another matter. . .'. It is unlikely that bacteria would be able either to carboxylate glutamate residues or to carry out the post-transcriptional glycosylation required to produce functional factor VIII, although production in yeast remains a possibility. We apologise for the error and thank Dr A.D. MacNicol for pointing it out.

Obituary

Albert Claude, 1899–1983

from Christian de Duve and George E. Palade

ALBERT CLAUDE, who died in Brussels on Sunday 22 May, belonged to that small group of truly exceptional individuals who, drawing almost exclusively on their own resources and following a vision far ahead of their time, opened single-handedly an entirely new field of scientific investigation.

He was born on 23 August 1899, in Longlier, a hamlet of some 800 inhabitants situated in the heart of the Belgian Ardennes. His mother developed breast cancer when he was 3 years old, and he was with her most of the time to witness the progress of the disease until she died four years later. He attended the village school for a few years, but then his family moved to the German-speaking village of Athus, where, as he has recounted, he found himself learning to read German in Gothic, without understanding it. Amazingly, his school education stopped there.

At the age of 12, he went to work in the local steel factory, where he had risen to the position of draughtsman by the time the First World War broke out. During the war, Claude served underground in occupied Belgium for the British Intelligence Service, and earned several military distinctions.

The activities and disturbances did not prevent Albert Claude from pursuing an intense process of self-education. His childhood dream had been to study medicine, but his lack of a high-school diploma barred his access to medical school. Reluctantly, he prepared, and passed successfully in 1921, the entrance examination to the School of Mining Engineering, for which a high-school diploma was not required. Then something of a miracle happened. He was able to take advantage of a government disposition — obviously not intended for him — allowing war veterans to enter a university without a diploma or examination. He immediately enrolled in the Medical School of the University of Liège — not without a good measure of apprehension as he believed that the courses were given in Latin. He graduated as an MD in 1928, one year ahead of the regular curriculum.

As a student, Claude was already fascinated by cells:

"I remember vividly my student days, spending hours at the light microscope, turning endlessly the micrometric screw and gazing at the blurred boundary which concealed from us the mysterious ground substance where, one felt, the secret mechanisms of cell life might be found. Until . . . I realized that I should stop that futile game, and should try something else. In the meantime, I had fallen in love with the shape and the color of the eosinophilic granules of leucocytes, and attempted to isolate them. I failed — and consoled myself later on in thinking that it was technically premature, especially for a premedical student, and that the eosinophilic granules were not pink, anyway."

(Out of Claude's Nobel lecture.)

Having failed in his first project, Claude went on to study the fate of the mouse sarcoma S-37 when grafted in rats. The thesis he wrote on his observations earned him a government scholarship which he used to go to Berlin. There he first worked at the Cancer Institute of the University, but was forced to leave prematurely after showing that the bacterial theory of cancer genesis propounded by Blumenthal, the Institute's director, rested on faulty experimental manipulations: the simultaneous inoculation of cancer cells with the incriminated bacteria. Claude then joined the Danish scientist Albert Fischer, a pioneer of tissue-culture techniques.

By then, Claude knew exactly what he wanted to do — isolate and characterize the agent of the Rous sarcoma — and where to do it — The Rockefeller Institute for Medical Research. With his characteristic mixture of naive directness and unselfconscious assurance, he wrote out a research project and sent it to Simon Flexner, the director of the Institute, asking to be admitted in one of the Institute's laboratories. It is to Flexner's credit that he reacted favourably to this unconventional approach and Claude sailed from Antwerp on Friday the 13th of September 1929 to spend the next 20 years at the Rockefeller Institute, first bringing to fruition his project on the Rous sarcoma virus, and moving on from there to prepare the fulfilment of his main dream: to enter the cell, 'the mansion of our birth'.

In 1949, he accepted a pressing offer from the Free University of Brussels to assume the directorship of the Jules Bordet Institute. Claude brought to his new duties the same thoroughness and perfectionist attention to detail that he had devoted to his scientific work. But it took him several years before he was able to return to the cell. He retired from the Bordet Institute in 1971, and moved to a new laboratory offered to him by the Catholic University of Louvain. In 1974, he was awarded the Nobel Prize in Physiology or Medicine. We were honoured to share it with him.

Claude's scientific career developed logically from a deep interest in cancer — the disease that killed his mother. His work at the Rockefeller Institute showed that the 'chicken tumour I agent' (Rous sarcoma virus) was a complex of ribonucleic acid, protein and phospholipid that lost its activity upon UV irradiation, the inactivation spectrum coinciding with the absorption spectrum of nucleic acids.

But then in the late 1930s, he discovered that similar complexes (without the biological activity of the agent) were present in large amounts in the cells of chick embryos he used as controls. These normal complexes became in time 'small granules' and finally

'microsomes'. That was the beginning of a rather long side voyage that led him into the heart of the cell, not in search of a virus, but with the determined intent to find out whatever there was in it in terms of isolatable particles, and to account for them in a careful quantitative manner. It was a glorious voyage during which he worked out his now classic cell-fractionation procedure. Most of what we know today about the chemistry and activities of subcellular components is based on his quantitative approach. Claude enjoyed isolating whatever was isolatable: from microsomes to 'large granules' (later recognized as mitochondria), chromatin threads and zymogen granules. The isolated particles were characterized in terms of their basic chemistry and — in the case of the large granules — in terms of enzymic activities in work done with G. Hogeboom, W. Schneider and R. Hotchkiss. It was a very impressive harvest for a relatively short period of about 8 years during which his attention was naturally and logically diverted to a new approach, electron microscopy.

In 1945 he succeeded with Keith Porter and Edward Fullam in getting electron micrographs of cultured fibroblasts in which 'a lace-like reticulum' could be clearly seen below the limit of resolution of the light microscope. In time, this reticulum became the now well known endoplasmic reticulum of all eukaryotic cells. Two years later, using essentially the same approach and working with K. R. Porter and E. Pickels, Claude finally found the chicken tumour I agent in infected cultured cells.

One year later, he gave a travelogue of that historic voyage at the Harvey Society. His lecture retraces in memorable fashion the construction of those two pillars of modern cell biology: cell fractionation and electron microscopy. It heralds the beginning of three decades of unprecedentedly rapid developments worked out in many laboratories throughout the world.

Albert Claude approached science and the other facets of human culture — he was a friend of the painter Diego Rivera and of the musician Varèse — with the candid and unprejudiced open-mindedness of the unschooled. In some ways, he could serve as a good example for the proponents of nature versus nurture. Yet, he was also a child of his environment, reflecting in his attitude the simple commonsense and fierce individualism of the Ardennais peasants, deriving his love of life and beauty from the rugged countryside in which he spent a lonely and dreamy childhood, marked by the "blues of the blueberries and of slate, the blue-green of the fir-trees, new comers among the oaks, the blue-grey of the covered skies, but also in full summer, by the clear waters and the black nights on the milky way" (quoted from a letter of Claude to Marcel Florkin). □

Christian de Duve is at the Catholic University of Louvain and at The Rockefeller University and George E. Palade is at the Yale University School of Medicine.

Lunar magnetism, polar displacements and primeval satellites in the Earth–Moon system

S. K. Runcorn

Institute of Lunar and Planetary Sciences, School of Physics, The University, Newcastle upon Tyne NE1 7RU, UK

The discovery that rocks of the lunar crust possess remanent magnetism is a surprising result in view of the fact that the present Moon does not exhibit a magnetic field, nor was there once evidence of an iron core in which one could be generated. This review looks at the concept of lunar palaeomagnetism and discusses its implications for lunar science as a whole.

THE discovery^{1–6} that the lava and breccia returned by the Apollo 11 mission from the Moon possessed remanent magnetization was one of the first, and perhaps the most unexpected, of the project and has become a controversial issue in lunar science. The consensus had been that the Moon today has no general magnetic field and that it has no iron core in which one could have been generated. Consequently the interpretation of this discovery has not been as straightforward—nor its significance as quickly appreciated—as the geochemical discoveries. The petrology of the lunar samples⁷ returned from the Apollo landings and the radioactive age determinations⁸ led to a quickly accepted view of the early history of the Moon; its differentiation, in which the anorthositic highland shell was formed, the early bombardment of the Moon by planetesimals, which produced the multi-ring basins, and the later flooding of those on the near side by basaltic lava generated by partial melting in the interior. In contrast, the explanations of lunar palaeomagnetism and of other phenomena of lunar physics, such as the non-hydrostatic gravitational anomalies, both the non-hydrostatic second harmonic⁹ and the mascons¹⁰, correlating with the lava flooded circular near-side maria, and the deep and shallow moonquakes¹¹, have remained controversial. This is not because lunar physics has less to tell us about the Moon's origin and evolution but because primeval physical events generally leave less diagnostic evidence than do primeval chemical processes.

Some years ago extensive reviews of palaeomagnetism appeared, giving careful accounts of the observations by satellite and surface magnetometers¹² and of laboratory studies¹³ on the returned samples but they did not reach any definite conclusions on the origin of the widespread magnetization of the crust. Many have concluded that 'we do not know if the Moon did once have a dipole field... we cannot draw on the remanent magnetization of the Moon to tell us about the structure of the interior'¹⁴. I now set out the historical development of the subject, and argue that such negative conclusions are not justified.

Remanent magnetism of the Moon

The existence of an iron core¹⁵ in the Moon was first suggested in an explanation of the non-hydrostatic second harmonic term in the Moon's figure. From Cassini's three laws of the Moon's rotation, Laplace had shown that this term is about 17 times greater than that expected on hydrostatic theory. It had been assumed that this bulge was a primeval distortion retained by the finite strength of the lunar interior—Jeffreys thought it was a fossil tidal bulge—but I suggested that it is maintained by solid state convection. To account for the second harmonic it was necessary to postulate a two-cell convection pattern, which suggested the presence of a core between 350 and 500 km in radius. By the convection theory it is possible to account for the difference between $(C-A)/C$, where C and A are the moments of inertia about the polar axis and the equatorial one

directed to the Earth respectively, and the actual ellipticity determined by astronomical methods: this difference is not explained on the previous theories of distortion of an original uniform body.

So when remanent magnetism was discovered in the samples of lava and breccia returned by Apollo 11 from Mare Tranquillitatis, we suggested that the Moon had once possessed a magnetic field generated by the dynamo process¹⁶ in an iron core. Soon afterwards, a magnetic field of 35 nT (3.5×10^{-4} G) was discovered at the Apollo 12 site and disturbances observed by the Explorer 35 magnetometer in the Moon's shadow, evidently caused by deflection of the solar wind by local magnetic anomalies on the Moon's limb, led to the discovery¹⁷ of small regions of the surface with magnetic fields of 10 nT coherent over many 10 km, termed 'magcons'¹⁷. This ubiquitous magnetism of the lunar surface was most simply explained by postulating the existence of an internally generated lunar magnetic field, a theory that was strengthened when the Apollo 12 lavas were also found to have stable remanent magnetization¹⁸ but alternatives were soon suggested. In the early development of terrestrial palaeomagnetism, the discovery of directions of magnetization in the Mesozoic, Palaeozoic and Precambrian different from those of Tertiary rocks, challenging widely held views about the fixity of continents, stimulated many attempts to explain the remanent magnetization of rocks by local causes, such as the strains associated with tectonics. So the discovery of lunar palaeomagnetism, challenging certain long held views about the Moon, such as that it does not possess an iron core¹⁹, stimulated suggestions that it arose from local processes, such as meteoritic²⁰ or cometary²¹ impacts, or electric current produced by contact potentials²² or even lightning²³ or electrical effects in volcanic eruptions²⁴; in all these explanations the magnetizing field was locally generated and transient. These necessarily vague theories, involving processes of which we have no adequate terrestrial analogues, made the formulation of decisive tests difficult. That some rocks on the Moon were magnetized by such processes, just as some terrestrial rocks are magnetized by lightning, is conceivable: but the fundamental question is whether the global phenomena of lunar magnetism can be explained by local processes unconnected with the fundamental evolution of the lunar interior.

The initial interpretation of the magcons was that they arise from crust magnetized on a similar scale which would support magnetization by impacts or other local processes. But if crustal magnetization is so restricted, was it not unlikely that unshocked rocks would have been brought back from all the Apollo mission sites possessing magnetization with some, though varying, stability? This paradox was resolved using a simple result of potential theory—that the field outside a uniformly magnetized infinite plate is zero. Thus it was concluded that the Moon's crust is generally magnetized and that the local magnetic fields arise from edge effects: crust demagnetized by impact or the large craters or contrasting magnetizations of adjacent strata

or deposits of different age or composition²⁵. A second paradox then arose: such a general magnetization of the crust would be expected to produce a small external dipole field but the present upper limit to this surface field had gradually been reduced from the first value of Luna 2 of 100 nT, until analysis of the Apollo 15 and 16 sub-satellite magnetometer observations gave 0.05 nT (ref. 26). The resolution of this paradox depends on a novel theorem^{27,28}. If a spherical shell becomes permanently magnetized in the same direction as and with an intensity proportional to a field of internal origin, that later disappears, the field outside the magnetized shell is zero, neglecting susceptibility. Thus because the present lunar dipole moment is not significantly different from zero, the explanation of lunar palaeomagnetism by the core dynamo theory is, paradoxically, supported. Of course, a randomly magnetized crust would also be consistent with the present zero moment and we will see later that the model is too simple. The possibility that the Moon's differentiated crust could have uniform magnetization acquired from cooling in an external field, for example, the geomagnetic field when the Moon was close to the Earth in its early history, is excluded for, by the converse of the above theorem, the Moon's crust would now have a small dipole moment.

Palaeointensities of lunar field

Scientific method requires the formulation of tests of the core dynamo theory and the 'local' theory. The first test emerged from the development of methods to determine the strength of the field responsible for the natural remanent magnetization (NRM). Such palaeointensity experiments compare the NRM with the remanent magnetization produced in a demagnetized rock sample by a laboratory magnetic field, simulating the natural process by which the NRM was acquired. Two methods are used in which step by step removal of the NRM and the step by step acquisition of a laboratory magnetization (thus sampling grains with different degrees of magnetic memory) are compared. In the Thellier-Thellier method the decrease in the NRM is observed as the rock is heated in zero field and the results compared with the thermoremanent magnetization (TRM) acquired in a small laboratory field as the sample is cooled from above the Curie point of the magnetic grains. In the lunar rocks, iron grains with a Curie point of 780 °C but different blocking temperatures carry the magnetization. The second method (the ARM method) is one in which the decrease in NRM under alternating field demagnetization is compared with the acquisition of an anhysteretic remanent magnetization (ARM) in a small d.c. field in the presence of gradually increasing alternating magnetic fields²⁹. These methods give consistent results only on those rocks with magnetic components that possess the requisite stability and many lunar rocks, like many terrestrial rocks, do not have this. We studied the experimental determinations from those rocks which gave self-consistent results and concluded³⁰ that the palaeointensity values fell near an exponential decay curve from the remarkably high palaeofield of 1 G at 4,000 Myr ago to a field of about 0.01 G at 3,200 Myr ago (Fig. 1). Although our results, especially the high intensity of the ancient field, were challenged, there now seems to be general agreement that such high fields were present 4,000 Myr ago on the Moon³¹, although some would seek explanations other than the core dynamo. This dependence of palaeointensity on age argues strongly that the magnetizing field was of internal origin, because while locally generated fields would not have had a simple dependence on age, a decrease in the field with decreasing age would be expected for a dynamo driven by internal energy sources that diminished with time. Some time later than 3,200 Myr ago the field disappeared either by the freezing of the core or the failure of the energy sources to maintain the magnetic Reynolds number above its critical value.

Recently an independent method of estimating palaeointensities has been developed depending on comparison of the NRM with the IRM (isothermal remanent magnetization)³².

While for a single sample this method is not as accurate as those described above, IRM information exists on many more Apollo samples. Cisowski³³, using this method, has obtained independent support for the existence of a palaeofield of ~1 G 3,900–3,800 Myr ago, subsequently decreasing rapidly. Before then he also finds some evidence for smaller palaeofields suggesting that the processes of field generation built up gradually. His strongest conclusion is that palaeointensity depends on age, supporting the hypothesis that the field is generated by a core dynamo.

Surveys of the data on palaeointensities have been presented by Sugiura and Strangway³⁴ and they plot a diagram (reproduced in ref. 35) that is very different from Fig. 1. Sugiura and Strangway only admit the validity of the Thellier-Thellier method, which has been done on relatively few lunar samples (because of possible chemical alteration on heating) so they omit most of the results in Fig. 1. But they include data on two high grade breccias giving low values of field ~0.01 G, which they plot at ~3,800 Myr (the age of the rock fragments in them) indicating this to be the maximum age. The times of magnetization are the ages at which the breccia were welded together, and cannot be determined by any present method but may be very much less than 3,800 Myr. They also include a young impact glass bead of 3 Myr which gives a palaeointensity³⁶ of 0.01 G. The glass bead is, however, a sample in which, *a fortiori*, the physical effect of impact may be a substantial factor in the magnetization process and is certainly not understood. Moreover, the palaeofield measured is only three times that of the highest crustal field measured by the Apollo 16 astronauts so that it may not be relevant to the discussion of the lunar dynamo. When the ARM method has been used on rocks on which the Thellier-Thellier method has also been used, there is agreement in the palaeointensities, so that there seems little reason to exclude from Fig. 1 the palaeointensities determined by the ARM method alone and this viewpoint has been substantiated by the recent work of Cisowski³³.

It has been suggested that the presence of solar wind gases in lunar breccias argues against the existence of a strong primeval dipole field, which would have excluded charged particles (except at the poles and possibly during reversals) but the age of the impregnation with solar wind is certainly younger than that of the rock fragments. Thus even breccias thought to be formed early may, in fact, have formed after the dipole field had disappeared.

Existence of a lunar core

The strong magnetic field, apparently generated by the core dynamo, raises very fundamental questions which did not arise in the geochemistry and petrology of the Moon. The fact that a lunar magnetic field existed 4,000 Myr ago, and as we will see probably about 4,200 Myr ago, requires there to have been an iron core at this time. At the beginning of the Apollo project, the discovery of the anorthositic composition of the highlands and the measurement by lunar seismology of its thickness on the near side, proved what had not seemed likely on the accretion theory of its origin, that at least the outer part of the Moon's interior had melted by 4,400 Myr. Geophysicists immediately proposed that the energy of the accreting bodies had been wholly trapped in the Moon and had heated up its outer part to the melting point, perhaps even forming a 'magma ocean'. But while the kinetic energy of a body falling with the present escape velocity, 2.3 km s⁻¹ is just equal to that required to melt it if it is wholly retained, the gravitational energy released by the accreting bodies would have been much smaller in the early stages of the Moon's accretion. Appreciating this, geochemists constructed models with a deep interior of primeval undifferentiated Solar System material, and were not slow in finding geochemical evidence in support. An iron core in this scenario was impossible.

Moreover only recently, long after the first suggestion of its existence¹⁵, has evidence for an iron core in the Moon today been found. Sonett, studying the rise of electrical conductivity

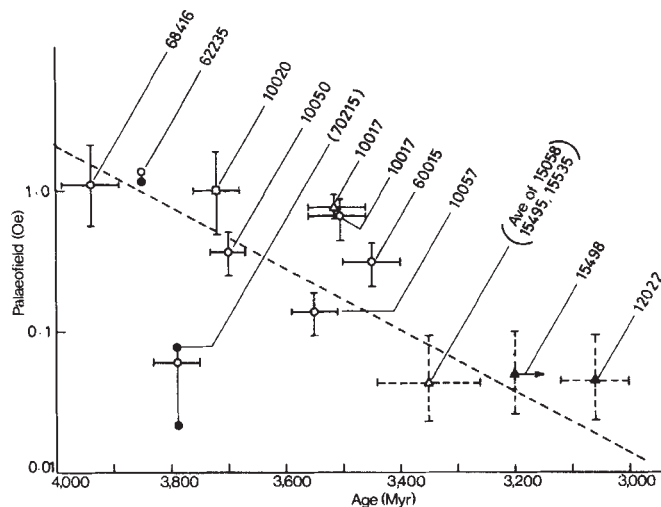


Fig. 1 Intensity of the ancient lunar magnetic field determined from laboratory experiments on Apollo samples.

in the deep interior from induction by solar wind variations, finds evidence³⁷ for a conducting core of radius not greater than 400 km, which could be an iron or iron-iron sulphide core. Russell *et al.*³⁸ determine the size of the core as the Moon passes through the oppositely directed quasi-steady fields of the Earth's magnetotail, finding a rather larger radius of 500 km. The P waves from one meteoritic impact on the far side received by the Apollo seismometer network gave evidence for an iron core of radius up to 350 km: the signal received by the Apollo seismometer was delayed³⁹ by 1 min compared with that expected for a wholly silicate Moon. The accepted value⁴⁰ for the moment of inertia factor (C/Ma^2) is 0.3905 ± 0.0003 , which, with the mean density of 3.34 and a simple equation of state⁴¹, gives an iron core up to 500 km in radius. None of these is conclusive evidence for an iron core; in each case other models can be generated to fit the observations, but there is now no reason to doubt its existence if it is required to explain lunar palaeomagnetism.

Palaeomagnetic directions

Recent determinations⁴²⁻⁴⁵ of the directions of the fields that magnetized different parts of the lunar crust lead to a second test of the opposed theories: do they fit a dipole field or are they random? Because the *in situ* orientations of the returned lunar samples are unknown, knowledge of the palaeomagnetic directions has had to await sophisticated analysis of the data obtained from the three component magnetometers in the Apollo 15 and 16 subsatellites, although only a limited area of the crust has been surveyed. The magnetic anomalies were at first fitted by dipoles, but it was found for the best fit their positions were 50–60 km below the surface. As the selenothermal gradient causes the Curie point to be exceeded at such depths, and as such models are equivalent to uniformly magnetized disks⁴⁶ at the surface some hundreds of kilometres in radius, more plausible sources of permanent magnetization are rock strata in the outermost crust. So an important inference was drawn: that these magnetic anomalies demand a high degree of uniformity of magnetization of these strata extending over hundreds of kilometres. This requirement presents no difficulty for the core dynamo theory but it has not been shown that it can be met by theories which depend on local field generation.

The first modelling of crustal magnetic anomalies was done on the lunar far side highlands by Coleman *et al.*, who concluded that the palaeomagnetic directions were apparently randomly distributed in the present equatorial plane. On the hypothesis that the magnetizing field was a dipole, I have calculated pole positions from these directions and found that the north magnetic poles lie in two groups antipodal to each other: the mean poles close to the east-west direction in the lunar equator⁴⁷.

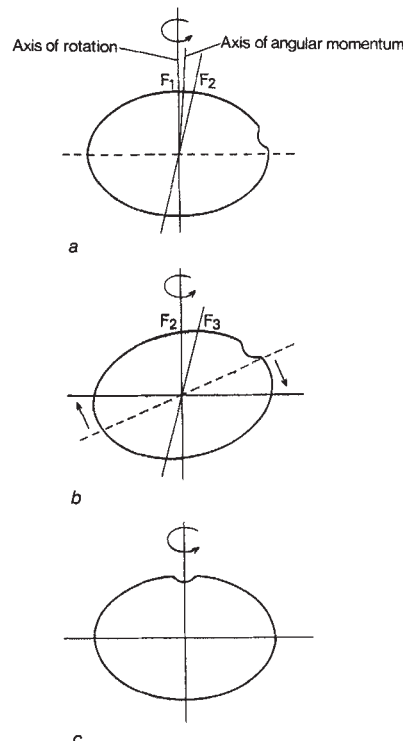


Fig. 2 The mechanism of polar wandering by impact basin formation. *a*, Just after impact; *b*, slow deformation of Moon by centrifugal force; *c*, final rotation state.

Hood⁴⁸ then discovered an anomaly over Reiner γ , an enigmatic light swirl-like feature on the lavas of Oceanus Procellarum. If it is a layer of ash or breccia, the absence of shadows shows that it cannot be thicker than 10 m. It cannot be the source as the intensity of magnetization would be too high. I have shown⁴⁹ that the pole position lay between the present pole and that of the far side sources. Recently Hood⁴⁹ had greatly extended the data by modelling with greater accuracy than those discussed above several new anomalies on the near side and a few on the far side. This outstanding contribution to selenophysics by Hood leads to a decisive test of the theories of lunar magnetism.

Hood⁴⁹ finds that these palaeomagnetic directions do not show the nonrandom planar distribution of the earlier data and concludes⁵⁰ that their apparent random distribution in space supports the magnetization by impact theory. In fact the idea is widespread in the literature that the magnetizations of the crust are chaotic or even random, (see, for example, Cook's account¹⁴ of the Moon). The surface field certainly changes direction on scales ranging from hundreds of kilometres to the short distances traversed by the astronauts. However, as I have shown above, uniform magnetization of a crustal layer does not give uniform field outside it. Magnetic fields are not like gravity fields which, as in the case of the mascons, have contours obviously revealing the uniform plates (of lava flooding the circular maria) which are their sources. It seems that the equation $\text{div } \mathbf{B} = 0$ is not sufficiently understood.

Reorientation of the Moon

Before going further with the interpretation of palaeomagnetic directions, we must consider whether the lunar magnetic dipole could have changed its orientation with respect to the Moon. As the Coriolis force would have been a dominant term in the magnetohydrodynamic equation of the Moon's core, the hypothesis that the mean dynamo field was a dipole directed along the axis of rotation is very securely based⁴⁶. However, good reasons exist for expecting that the Moon could have been reoriented with respect to its axis of rotation during the period of basin formation and their later flooding by lava forming the

Table 1 Mean pole positions

	Lat. (°N)	Long. (°E)	Radius of circle of confidence (deg)	Axis of rotation from mean poles	
	Lat. (°N)	Long. (°E)		Lat. (°N)	Long. (°E)
Imbrian–Upper Nectarian poles					
Sites from Oceanus Procellarum, Cayley formation and south of Mare Crisium					
Based on disks	–48.4	340.8	23.3	43.3	169.7
Based on anomalies	–60.2	339.2	31.6		
Based on disks	37.8	177.3	7.7		
Based on anomalies	40.4	182.4	17.1		
Lower Nectarian poles					
Sites on central and eastern near side and north-east far side					
Based on disks	49.6	338.3	19.6	48.9	154.4
Based on anomalies	54.9	335.8	29.4		
Based on disks	–51.2	165.9	56.2		
Based on anomalies	–48.5	178.0	60.2		
Pre-Nectarian poles					
Sites from western and south central far side of Moon					
Based on disks	–2.9	275.1	53.2	8.5	90.4
Based on anomalies	–7.2	274.6	54.4		
Based on disks	12.7	89.5	42.6		
Based on anomalies	22.6	92.0	59.8		
Sources on far side earlier analysis (based on 35 anomalies)					
	–2.4	93.9	34.3	1.0	92.8
	–4.3	271.7	19.7		

mascons. So tests of the dipole hypothesis of the origin of crustal remanent magnetism are not straightforward.

I have suggested that the Moon had been reorientated with respect to the axis of rotation through 90° and by about 50° since the times of magnetization of the far side highlands⁴⁷ and Reiner γ (ref. 46) respectively. Such ideas of 'polar wandering' as opposed to permanency or stability of the pole of rotation with respect to a planet's surface were once discussed^{51,52} as a possible explanation for palaeoclimatic anomalies in the Earth's geological record and for the 'polar wandering path' obtained from the palaeomagnetic data from the geological column in Great Britain before data from other continents gave the different paths which could only be reconciled on the continental drift hypothesis. The dynamical principle of these polar displacements is clear: if a redistribution of mass on the rigid lithosphere is brought about, and if flow in the interior is possible through solid state creep, considerable polar displacement can occur whereas if the interior is rigid, or if transient creep only occurs, the pole is only displaced by a small angle depending on the ratio of the redistributed mass to the equatorial bulge. In the early Moon impacts excavated deeply into the less dense anorthosite shell, 40 km thick on the near side, even after the reestablishment of isostatic equilibrium by the upwarping of the denser mantle would have left a basin ~10 km deep. Suppose a basin had been formed in low latitude the axis of maximum moment of inertia would have been displaced from F1 to F2 (Fig. 2a) and the three axes, of figure, of angular momentum and of instantaneous rotation, would have executed an eulerian nutation (compare with the Earth's Chandlerian wobble). After the wobble had been damped out, the Moon would have rotated, by Euler's theorem, about its new axis of figure (F2)—with the hydrostatic equatorial bulge no longer exactly in the plane perpendicular to it. The stress differences resulting from the centrifugal forces would have caused flow, assuming solid state creep beneath the lithosphere, endeavouring to return the hydrostatic equatorial bulge to the plane exactly perpendicular to the new axis of rotation (Fig. 2b). The axis of figure then moved to F3 nearer the basin and the process proceeded, in the absence of other basin forming events, until the basin was at the pole (Fig. 2c). The time scale of this process depends on the viscosity of the Moon's interior and on the properties of the lithosphere, the thickness of which at the time would be less than today. We do not know these para-

meters so that a calculation is probably less useful than to note that the ellipticity of the Earth has changed in response to its decelerating spin with a time constant not more than a few million years.

Palaeopoles

Because of the likelihood of the Moon's reorientation with respect to the axis of rotation, the test of the hypotheses concerning the nature of the magnetizing field cannot be done by comparing pole positions for all the directions determined: and the first essential step is to compare directions of magnetization only of those crustal strata magnetized in a time less than the time scale of polar wandering. If the directions are random in space the local field hypothesis would be proved: if they fit a dipole field the core dynamo hypothesis would be proved. However, we know too little about lunar geology and the nature of the sources to proceed by such a direct route. Instead I have determined⁵³ pole positions from Hood's new directions of magnetization and found empirically that the poles fall into three groups: each group aligned along three different axes with some poles antipodal to the other poles, the mean poles of each opposed group being roughly at 180°. This antipodal arrangement disposes of the view that the poles are random. It is, in fact, of fundamental significance: its explanation is that the lunar field reversed its polarity, as the geomagnetic core dynamo has done frequently, or that the intensities of magnetization or thicknesses of the magnetized rock strata vary, so that while the sources approximate to disks, some have magnetization parallel and some anti-parallel to the original magnetizing field. That the antipodal relation is statistically significant can be seen in Table 1 and in Fig. 3a, b and strongly suggests a global rather than a local origin for the field.

I have identified⁵³ the ages of the magnetization of the three groups of sources by the following arguments. First, those poles of Fig. 3b, derived from sources on Oceanus Procellarum (RG, KP, C, E), cannot quantitatively be explained by the lavas because they are too thin and their remanent magnetization too small. In the case of RG the correlation between the magnetic anomaly and the light-coloured feature has been much debated. However, even if it is a superficial deposit it is so thin that it is even less plausible as a source of the anomaly than the underlying lavas. The most reasonable explanation seems to be that the magnetic field has deflected solar wind particles

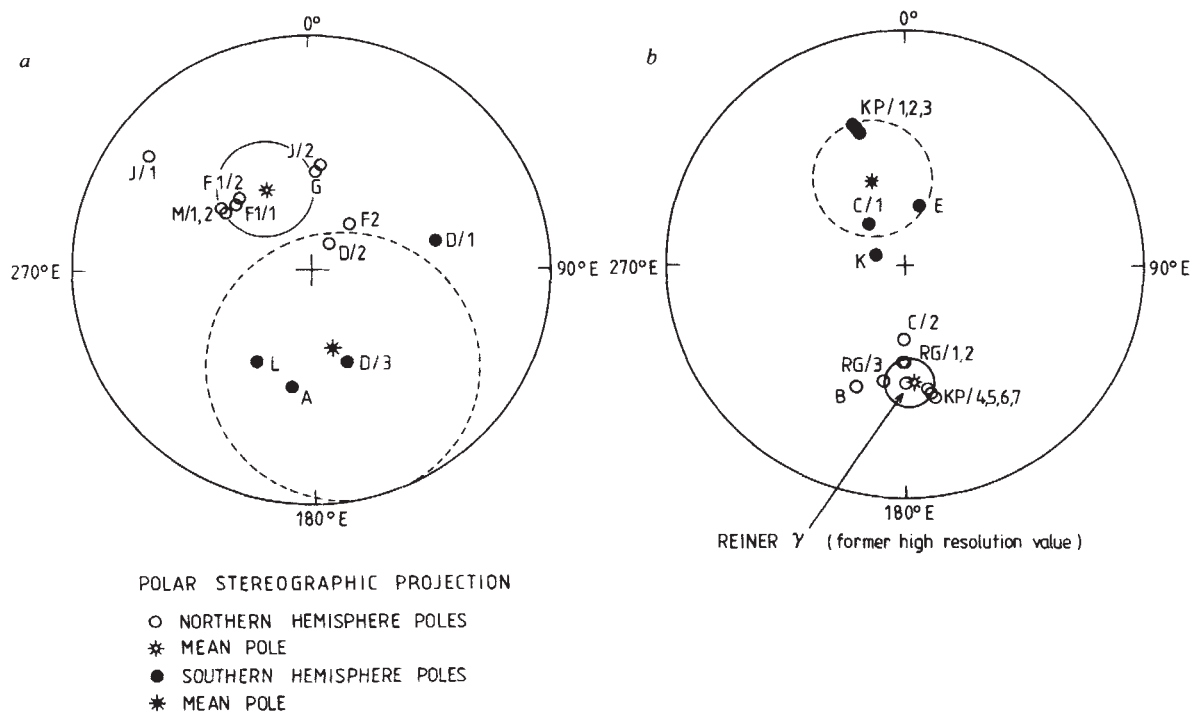
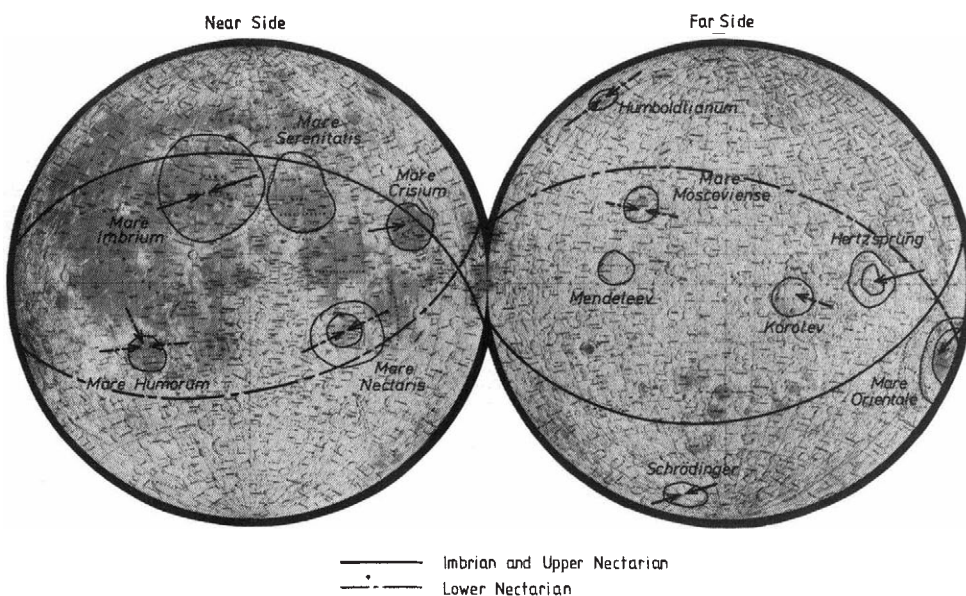


Fig. 3 *a*, Lower Nectarian pole positions from sources on central and eastern near side and north-east far side. *b*, Imbrian and Upper Nectarian pole positions from sources on Oceanus Procellarum, the Cayley formation and south of Mare Crisium. Means and circles of confidence calculated treating each disk separately.

Fig. 4 Palaeoequators: dotted line, Imbrian and Upper Nectarian; solid line, Lower Nectarian corresponding to the mean poles of Fig. 3*b* and *a*, respectively. Positions of sources of the modelled magnetic anomalies. Lettered sources of this and Fig. 5 are those modelled by Hood⁵⁰. Arrows show direction or possible alternative directions from which the impacting body came.



away from that part of the regolith so that the mineral grains have not been darkened by charged particle bombardment. It, therefore, follows that these anomalies arise from the underlying Fra Mauro formation. Source B is on the Cayley formation. Both formations are ejecta from the Imbrium impact so that the age of these poles is that of the Imbrium impact, which is 3,850 Myr. Source K is on the ejecta from the Crisium impact of Upper Nectarian age so this pole is 3,900 Myr in age. I then argued by analogy that the poles of Fig. 3*a* are of Lower Nectarian age, as J, M are on the Jansen formation, which is ejecta from the Nectarian impact of ~4,000 Myr and F1, F2 on the ejecta of Mendeleev also of Lower Nectarian age. I have also shown⁵³ that sources F, G, VDG, GER which also give bipolar, though very scattered, groups, have means not significantly different from the mean poles of the original far side sources. They are Preneptarian in age, 4,200 Myr. As the Preneptarian multi-ring basins and the magnetization ages span

a longer interval of time, being the oldest period, the considerable scatter in the poles is to be expected. Anderson and Wilhelms⁵⁴ had concluded that a correlation exists between the strongest magnetic anomalies on the far side and the ringed impact basins and they attributed these anomalies to thick deposits of basin ejecta. Therefore, I conclude that these magnetizations are acquired by the basin ejecta after emplacement either through cooling from about the Curie point of iron as a TRM or as a CRM resulting from the formation of iron grains by reduction of silicate minerals.

The sources of the magnetic anomalies are plotted in Fig. 5. As the sources of Fig. 3*b* are securely dated as nearly contemporaneous, these data lead to a decisive test of the hypothesis of a primeval dynamo field. Figure 6 shows their angles of magnetic inclination plotted against their angular distances from the magnetic north pole at the time: the agreement with the dipole relation is good. Also shown is the relationship expected

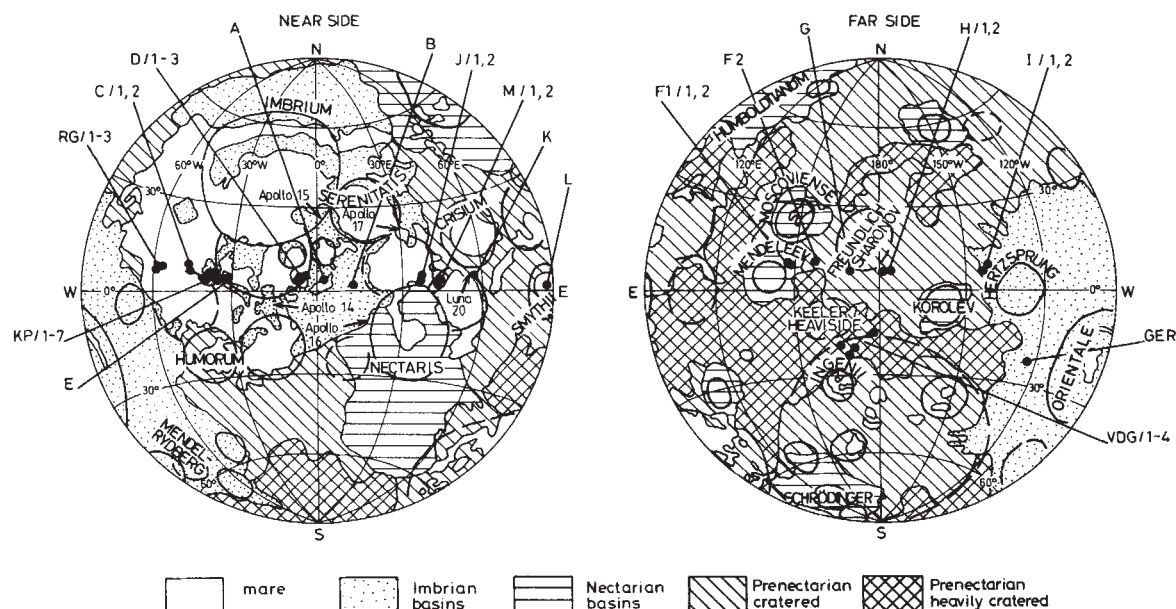


Fig. 5 Generalized geological map from ref. 35. Lettered sources of this and Fig. 4 are those modelled by Hood⁵⁰.

were the lunar crust magnetized by the Earth at the time of close approach or by another external uniform field. Both relationships are unaffected by field reversals on this plot, so that the crustal magnetization data prove the existence of an early lunar field of internal origin, unless it can be shown that the age assignments are entirely wrong.

Palaeoequators

A further argument that these groupings of poles are not an artefact is that their palaeoequators, calculated as in Table 1 from the means of the opposed mean poles, have a striking relationship with the multi-ring basins. The Imbrian–Upper Nectarian equator lies close⁵³ to the basins of that age, Orientale, Imbrian, Serenitatis, Crisium and on the far side Herzprung and Schrödinger as determined by crater counting⁵⁵. Similarly the Lower Nectarian equator places the basins of that age⁵⁵ in low latitude; Humorum and Nectaris on the near side and Mendeliev, Moscoviense, Korolev and Humboldtianum and Mendel–Rydberg on the far side. These relationships are shown in Fig. 4. I have also shown⁵³ that the Preneectarian equator has a similar relationship to basins of this age: only 3 out of 12 listed by Wilhelms⁵⁵ as certain basins are more than about 20° from it.

These associations between palaeoequators and impact basins show that the hypothetical mechanism discussed above for successive reorientations of the Moon with respect to its axis of rotation actually occurred. However, the evidence shows that at three epochs several impact basins were formed over times such that little polar displacement occurred. In such a case the subsequent motion of the pole would have been to a point on the great circle joining the basins rather than to one basin but through 90°. Successive displacements of the pole through about 90° do seem to have occurred as shown in Table 1.

Primeval satellites of the Moon

A further inference can be drawn from the observation that the multi-ring basins were formed in low latitude. The bodies which produced them could not have been in heliocentric orbits, originating in the asteroid belt or beyond: the chance of such impacts being always in the equatorial region is negligible. Consequently I have concluded that they were satellites in the Earth–Moon system⁵³. Were they Earth satellites along with the Moon, their geocentric orbits would have been nearly coplanar with the Earth's Equator and they would all have

retreated from the Earth by tidal friction at rates proportional to their masses, but would have been overtaken by the Moon. Even if these orbits were exactly coplanar the Sun's attraction would have caused them to precess at different rates depending on their distance from the Earth and to arrange that they impacted the Moon always near its equator is too contrived. Thus I conclude that they were satellites of the Moon, which would in the presence of damping mechanisms, such as rings of dust, settle into orbits exactly coplanar with the Moon's equatorial bulge, which then was a hydrostatic one⁵⁶. Their orbits would have decayed by tidal friction and they would have impacted the Moon at glancing angles. Laboratory experiments⁵⁷ on hypervelocity impacts and studies of missile impacts⁵⁹ show that in impacts at <5° debris is preferentially ejected at right angles to the path—the bowtie or butterfly effect—and elongation of the crater and a narrower spurt of ejecta in the forward direction occur. I have pointed out⁴⁶ that the geological studies of Mare Orientale show these effects consistent with a body in orbit parallel to the corresponding palaeoequator as in Fig. 4. McCauley *et al.*⁵⁹ showed that the impacting body came from the north-east.

Recently, Wilhelms⁶⁰ has examined the earlier multi-ring basins and though some of the features are obscured by later blankets of ejecta, he is able to determine the directions (or alternative directions) of the incoming bodies. These are plotted in Fig. 6 and in almost every case are in good agreement with the conclusion reached above that the satellites were in decaying orbits exactly coplanar to the palaeoequator at the time. From these studies of the asymmetries of the basins in three cases out of the four where the sense as well as the direction can be determined, the satellites were in retrograde orbit: tidal friction would have pulled them into the Moon. Their velocity at impact would have been ~3 km s⁻¹ rather than the 10–15 km s⁻¹ of a body in heliocentric orbit: no diagnostic evidence exists as to whether the impacts producing the basins were low or high velocity ones, but less melted rock was found in the multi-ring basin ejecta than seems likely for the high-velocity impact.

Energy sources in the primeval Moon

What were the heat sources between 4,550 and 4,400 Myr ago that melted the Moon forming the iron core and what was the energy source driving the dynamo with gradually diminishing vigour to at least 3,200 Myr ago? Libby, Libby and I^{61–63} have argued that an exotic source, such as superheavy elements⁶⁴, existed in the early Solar System melted the Moon and as these

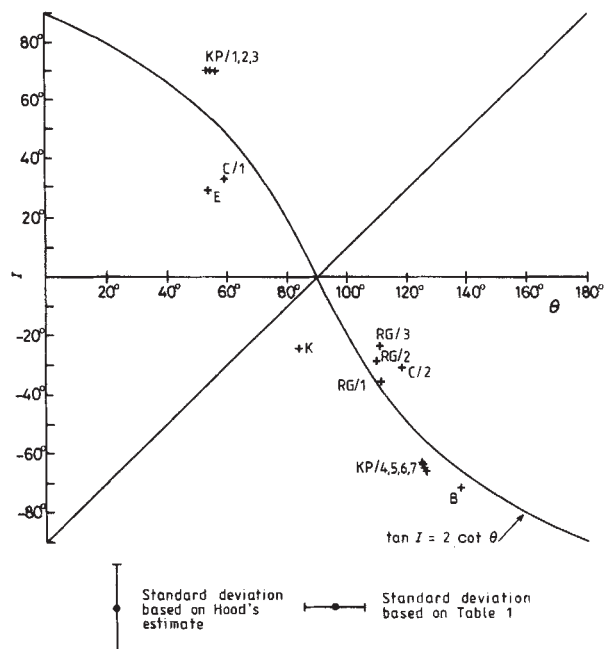


Fig. 6 Observed angle of magnetic inclination (I) against the angular distance (θ) to the mean north pole from data of Fig. 3b for Imbrian and Upper Nectarian age. Curved line shows relation of angle of magnetic inclination with co-latitude for an axial dipole field of either sign.

elements of atomic number 114 are siderophile or chalcophile they would have concentrated in the core causing convection and generating a field by dynamo action: the early melting and the rapid decay of the field being both explained by the relatively short half life (100 Myr) of these hypothetical elements. We had rejected accretional energies as insufficient to melt the Moon, and also Al^{26} as having far too short a half life to run the dynamo and, of course, aluminium is not siderophile. These speculations⁶⁵ have stimulated examination of the superheavy element hypothesis by searching for the likely fission products in primitive material of the Solar System, that is, in iron meteorites, or by searching for fission tracks in silicate inclusions in iron meteorites. So far the results⁶⁵⁻⁶⁷ have been negative, but not, we argue, conclusively so⁶⁸. In particular, our argument⁶⁹ that bimodal distribution of abundances of elements in iron meteorites might be fission products from superheavy elements in the early Solar System is wrong⁶⁵ because it now is clear that the early compilations of data on iron meteorites we used were unreliable because removal of silicate inclusions was not then adequately carried out. The nature of the heat sources in the early Moon therefore remains an unsolved and intriguing problem.

Conclusions

I conclude that the reorientations of the Moon with respect to its axis of rotation took place between 4,200 and 4,000 Myr ago through the formation of the Pre-nectarian basins, between 4,000 and 3,900 Myr ago due to Lower Nectarian basins and between 3,900 and 3,850 Myr ago due to Upper Nectarian and Imbrian basins. Much later between 3,600 and 3,200 Myr ago the circular basins on the near side were flooded by lava⁷⁰ forming the mascon gravitational anomalies and after the solidification of the lava the present non-hydrostatic second degree harmonic in the lunar figure was produced by solid state convection with a two-cell pattern. These processes stabilized the pole in its present position: Melosh^{71,72} has calculated for a model of the Moon (a uniform sphere with surface masses simulating the mascons), that the axes of maximum and minimum moments of inertia lie within 5° – 10° of the present axis of rotation and the direction to the Earth at nodal passage respectively. Note that had the mascons or the present second harmonic in the figure been present 4,200–3,800 Myr ago the pole would not have moved. The complex problems in dating these events are discussed by Turner⁷⁴.

The possibility that satellites of the inner planets have been lost by orbital decay has been discussed^{74,75}. That there could have been satellites around the early Moon was raised⁷⁶ but their lifetimes were disputed^{77,78}. The evidence presented here for lunar satellites in the early history of the Earth–Moon system may be an important clue concerning its origin.

Possible correlation between the light swirl features and strong magnetic anomalies have been held to favour the crustal magnetization being produced by impact: the swirls being attributed to nearby meteorite impacts⁷⁹ or comets⁸⁰ and given recent ages. Geophysical correlations are notoriously misleading and a possible interpretation has been suggested for the correlation: the swirls are areas of the regolith which have not been darkened because solar particles have been deflected by the strong local magnetic fields⁸¹. However, these strongly magnetized regions, mapped by the reflected electron technique^{82,83}, are antipodal to the impact basins⁸⁴. It seems possible that the shock waves from such great impacts, focusing 180° away, cause precipitation of iron particles from the silicates which then acquire a strong CRM from the lunar field. Such a process seems required to understand the early attempts^{85,86} to correlate geology and magnetism in which strong magnetic anomalies correlated better with ejecta blankets than with other lunar features.

I thank all those lunar scientists to whose work I have referred, especially Drs S. M. Cisowski, P. J. Coleman, D. W. Collinson, M. Fuller, L. L. Hood, R. P. Lin, C. P. Sonett, A. Stephenson, D. W. Strangway and D. E. Wilhelm for many discussions on lunar magnetism. Some of this paper was written while I was a visiting professor in the University of Queensland, Australia and Hitchcock Visiting Professor in the University of California, Berkeley. I thank Dr L. Molyneux, Miss M. Hopkinson and Mrs D. Cooper for help.

1. Runcorn, S. K. *et al. Geochim. cosmochim. Acta Suppl.* **1**, 2369–2387 (1970).
2. Doell, R. R., Gromme, C. S., Thorpe, A. M. & Senftle, F. E. *Geochim. cosmochim. Acta Suppl.* **1**, 2097 (1970).
3. Helsey, C. E. *Geochim. cosmochim. Acta Suppl.* **1**, 2213 (1970).
4. Larochelle, A. & Schwarz, E. J. *Geochim. cosmochim. Acta Suppl.* **1**, 2305 (1970).
5. Nagata, T. *et al. Geochim. cosmochim. Acta Suppl.* **1**, 2325 (1970).
6. Strangway, D. W., Larson, E. E. & Pearce, G. W. *Geochim. cosmochim. Acta Suppl.* **1**, 2435 (1970).
7. Taylor, S. R. *Lunar Science: A Post Apollo View* (Pergamon, Oxford, 1975).
8. Wasserburg, G. J., Papanastassiou, D. A., Tera, F. & Huneke, J. C. *Phil. Trans. R. Soc. A* **285**, 7–22 (1977).
9. Runcorn, S. K. *Proc. R. Soc. A* **296**, 270–284 (1967).
10. Muller, P. M. & Sjogren, W. L. *Science* **161**, 680–684 (1968).
11. Lammlein, D. R. *Phil. Trans. R. Soc. A* **285**, 451–461 (1977).
12. Dyal, P., Parkin, C. W. & Daily, W. D. *Rev. Geophys. Space Phys.* **12**, 568–591 (1974).
13. Fuller, M. *Rev. Geophys. Space Phys.* **12**, 23–70 (1974).
14. Cook, A. H. *Interiors of the Planets* (Cambridge University Press, 1980).
15. Runcorn, S. K. *Nature* **195**, 1150–1151 (1962).
16. Runcorn, S. K. *et al. Science* **167**, 697–699 (1970).
17. Mihalov, J. D., Sonett, C. P., Binbaik, J. H. & Moutsoulas, M. D. *Science* **171**, 892–895 (1971).
18. Runcorn, S. K. *et al. Proc. R. Soc. A* **325**, 157–174 (1971).
19. Cook, A. H. *The Physics of the Earth and the Planets* (Macmillan, London, 1973).
20. Hide, R. *The Moon* **4**, 39 (1972).
21. Gold, T. & Sotor, S. *Planet. Space Sci.* **24**, 45 (1976).
22. Dyal, P., Parkin, C. W. & Daily, W. D. *Geochim. cosmochim. Acta Suppl.* **8**, 767–784 (1977).
23. Nagata, T., Ishikawa, Y., Kinoshita, H., Kono, M., Syono, Y. & Fisher, R. M. *Geochim. cosmochim. Acta Suppl.* **1**, 2325–2340 (1970).
24. Cap, F. F. *J. geophys. Res.* **77**, 3328–3333 (1972).
25. Runcorn, S. K. & Sonett, C. P. *The Moon* **8**, 308–334 (1973).
26. Russell, C. T., Coleman, P. J. & Schubert, G. *Science* **186**, 825–826 (1974).
27. Runcorn, S. K. *Nature* **253**, 701–703 (1975).
28. Runcorn, S. K. *Phys. Earth planet. Inter.* **10**, 327–335 (1975).
29. Stephenson, A. & Collinson, D. W. *Earth planet. Sci. Lett.* **4**, 361–369 (1974).
30. Stephenson, A., Runcorn, S. K. & Collinson, D. W. *Geochim. cosmochim. Acta Suppl.* **6**, 3049–3062 (1975).
31. Sugiura, N. & Strangway, D. W. *Geochim. cosmochim. Acta Suppl.* **14**, 1801 (1980).
32. Cisowski, S. M., Fuller, M. D., Wu, Yee Ming, Rose, M. F. & Wasilewski, P. J. *Geochim. cosmochim. Acta Suppl.* **6**, 3123 (1975).
33. Cisowski, S. M. *Adv. Space Res.* **2**, 31–34 (1983).
34. Sugiura, N. & Strangway, D. W. *J. geophys. Res. Suppl.* **88**, A684–A690 (1983).
35. Taylor, S. R. *Planetary Science: a Lunar Perspective* (Lunar and Planetary Institute, Houston, 1982).
36. Sugiura, N. & Strangway, D. W. *Geochim. cosmochim. Acta Suppl.* **10**, 2189 (1979).
37. Sonett, C. P. *Rev. Geophys. Space Phys.* **20**, 411–455 (1982).

38. Russell, C. T., Coleman, P. J. & Goldstein, B. E. *Geochim. cosmochim. Acta Suppl.* **12B**, 831-836 (1981).
39. Nakamura, Y. *et al. Geophys. Res. Lett.* **1**, 137-140 (1974).
40. Ferrari, A. J., Sinclair, W. S., Sjogren, W. L., Williams, J. G. & Yoder, C. F. *J. geophys. Res.* **85**, 3939-3951 (1980).
41. Baumgardner, J. R. & Anderson, O. L. *Adv. Space Res.* **1**, 159-176 (1981).
42. Hood, L. L., Russell, C. T. & Coleman, P. J. *Geochim. cosmochim. Acta Suppl.* **9**, 3057-3078 (1978).
43. Hood, L. L., Coleman, P. J. & Wilhelms, D. F. *Science* **204**, 53-57 (1979).
44. Hood, L. L., Russell, C. T. & Coleman, P. J. *J. geophys. Res. Lett.* **5**, 305-308 (1978).
45. Hood, L. L., Russell, C. T. & Coleman, P. J. *J. geophys. Res.* **86**, 1055-1069 (1981).
46. Runcorn, S. K. *Geochim. cosmochim. Acta Suppl.* **11**, 1867-1877 (1980).
47. Runcorn, S. K. *Nature* **275**, 430-432 (1978).
48. Hood, L. L. *Geochim. cosmochim. Acta Suppl.* **11**, 1879-1896 (1980).
49. Hood, L. L. *Geochim. cosmochim. Acta Suppl.* **12B**, 817-830 (1981).
50. Hood, L. L. *EOS* **62**, 161-163 (1981).
51. Gold, T. *Nature* **175**, 526-529 (1955).
52. Runcorn, S. K. *Phil. Mag. Suppl.* **4**, 244-291 (1955).
53. Runcorn, S. K. *Phys. Earth Planet. Inter.* **29**, 135-147 (1982).
54. Anderson, K. A. & Wilhelms, D. E. *Earth planet. Sci. Lett.* **46**, 107-112 (1979).
55. Wilhelms, D. E. *Rep. Planetary Geology Prog.* 1978-79 (eds Boyce, J. & Collins, P. S.), (NASA, Tech. Mem 80339, 1979).
56. Runcorn, S. K. *Phil. Trans. R. Soc. A* **385**, 507-516 (1977).
57. Gault, D. E. & Wedekind, J. A. *Geochim. cosmochim. Acta Suppl.* **9**, 3843-3875 (1978).
58. Moore, H. J. *Missile Impact Craters—White Sands Missile Range New Mexico—Applications to Lunar Research* (USGS Prof. Pap. No. 812C, 1976).
59. Scott, D. H., McCauley, J. F. & West, M. N. *Geologic Map of the West Side of the Moon* (US Geological Survey, Map I-1034, 1977).
60. Wilhelms, D. E. *Geologic history of the Moon* (USGS Prof. Pap., in the press).
61. Runcorn, S. K., Libby, L. M. & Libby, W. F. *Nature* **270**, 676-681 (1977).
62. Runcorn, S. K. *Science* **199**, 771-773 (1978).
63. Runcorn, S. K. *Earth planet. Sci. Lett.* **39**, 193-198 (1978).
64. Herrman, G. *Nature* **280**, 543-549 (1979).
65. Nozette, S. & Boynton, W. V. *Science* **214**, 331-333 (1981).
66. Bull, R. K. *Nature* **282**, 393-394 (1979).
67. Kaiser, T., Piepgras, D. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **52**, 239-250 (1981).
68. Runcorn, S. K., Libby, W. F. & Libby, L. M. *Nature* **287**, 565 (1980).
69. Libby, L. M., Runcorn, S. K. & Libby, W. F. *Nature* **278**, 613-617 (1978).
70. Runcorn, S. K. *Geochim. cosmochim. Acta Suppl.* **5**, 3115-3126 (1974).
71. Melosh, H. J. *Earth planet. Sci. Lett.* **25**, 322-326 (1975a).
72. Melosh, H. J. *Earth planet. Sci. Lett.* **26**, 353-360 (1975b).
73. Turner, G. *Phys. Chem. Earth.* **10**, 145-195 (1977).
74. Burns, J. A. *Nature phys. Sci.* **242**, 23-24 (1973).
75. Ward, W. R. & Reid, M. J. *Mon. Not. R. astr. Soc.* **164**, 21-32 (1973).
76. Reid, M. J. *Icarus* **20**, 240-248 (1973).
77. Gold, T. *Icarus* **24**, 134-135 (1975).
78. Reid, M. J. *Icarus* **24**, 136-138 (1975).
79. Schultz, P. H. & Srnka, L. J. *Nature* **284**, 22-26 (1980).
80. Srnka, L. J., Martelli, G., Newton, G., Cisowski, S. M. & Fuller, M. *Earth planet Sci. Lett.* **42**, 127-137 (1979).
81. Hood, L. L. & Schubert, G. *Science* **208**, 49-51 (1980).
82. Lin, R. P. *et al. Geochim. cosmochim. Acta Suppl.* **7**, 2691-2703 (1976).
83. Lin, R. P. *Phys. Earth planet. Inter.* **20**, 271-280 (1979).
84. Lin, R. P. *et al. Lunar planet. Sci.* **11**, 626-627 (1980).
85. Russell, C. T. *et al. Geochim. cosmochim. Acta Suppl.* **8**, 1171-1185 (1977).
86. Strangway, D. W., Rylaarsdam, J. C. & Annan, A. P. *Geochim. cosmochim. Acta Suppl.* **6**, 2974-2975 (1975).

ARTICLES

Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes

Hartmut Land, Luis F. Parada & Robert A. Weinberg

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, and Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02139, USA

Transfection of embryo fibroblasts by a human ras oncogene does not convert them into tumour cells unless the fibroblasts are established and immortalized before transfection. The embryo fibroblasts become tumorigenic if a second oncogene such as a viral or cellular myc gene or the gene for the polyoma large-T antigen is introduced together with the ras gene.

A LARGE and varied body of evidence indicates that carcinogenesis is a process involving multiple, independent steps. Epidemiological studies have suggested that cancer arises in proportion to a multiple power of elapsed lifetime. Pathological studies show that tumours progressively acquire new phenotypes by passing through a series of distinct stages such as anaplasia, metaplasia and neoplasia. Moreover, in many systems, experimental induction of a tumour requires at least two distinct types of stimulus, such as an initiator and a promoter (for example, see refs 1-5).

The evidence seems to be more equivocal at the cellular level. Although primary cultures of rodent cells can apparently become tumorigenic in a single step after infection *in vitro* by tumour viruses such as polyoma and adenovirus, this departure from multi-step carcinogenesis is more apparent than real. Recent studies have shown that each virus carries at least two genes encoding distinct functions, both of which must be expressed in order to realize the tumour cell phenotype⁶⁻⁸. Such work suggests that multi-step carcinogenesis might have an explanation at the genetic level: each step may require the activation of a distinct gene and the final phenotype may require the concomitant expression of many of the previously activated genes.

Other precedents support a model of multiple, cooperating, independently-activated genes. Induction of bursal lymphomas

by avian leukosis virus seems to require the activation of two separate oncogenes during lymphomagenesis. The *myc* gene becomes activated by adjacent insertion of a provirus^{9,10}, while the *B-lym* gene acquires activity via a second, distinct mechanism whose nature is unclear¹¹. This theme has been echoed in our own laboratory in a study of a promyelocytic leukaemia and an American Burkitt's lymphoma: in each case, the tumour cells carry altered versions of the *myc* gene, as well as activated versions of a second cellular oncogene, termed *N-ras*¹².

In apparent exception to this model of multiple genetic alteration, other studies have shown that a single oncogene can impart morphological alteration and tumorigenicity to NIH 3T3 mouse fibroblasts. The oncogene is usually introduced into the NIH 3T3 cells via calcium phosphate-mediated DNA transfection, which often results in establishment of multiple copies of the oncogene in the recipient fibroblasts¹³ although a single copy of the gene is enough to produce full transformation (C. Tabin and R.A.W., unpublished results).

One possible explanation of this paradox is that NIH 3T3 cells, which were chosen because of their particular competence in taking up and expressing exogenous DNAs¹⁴, behave abnormally in their response to oncogenes. These NIH 3T3 cells had been established (that is, adapted to grow indefinitely in monolayer culture) and then passaged extensively *in vitro*¹⁵. Thus, it seemed likely that they would deviate substantially

Table 1 Transformation of Rat embryo fibroblasts and Rat-1 cells following transfection of the EJ c-Ha-ras-1 oncogene

Cells	Transfected DNA	Colonies or foci per 10 ⁶ transfected cells				Tumorigenicity in nude mice (no. of tumours/no. of injections)
		No of foci normal medium	Ecogpt selection medium monolayer		Normal medium soft agar	
			No. of colonies	% Of colonies with morphologically transformed cells		
3° REF	pSV2gpt	0	150	0	0	0/10
3° REF	pEJ6.6+pSV2gpt	0	200	80-90	200	0/11
Rat-1	pSV2gpt	0	1,200	0	0	0/5
Rat-1	pEJ6.6+pSV2gpt	2,400	1,200	40-50	2,400	6/6

Primary cultures of REFs were prepared as described elsewhere⁶⁰ from 12–14-day-old Fisher rat embryos; 3–4 days later the cells were passaged and 1.2×10^6 cells were seeded onto 100 mm Petri dishes in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Hyclone)(normal medium). Parallel cultures of the Rat-1 cell line²⁹ were seeded at a density of 5×10^5 cells (normal medium) so that 18–24 h later both types of culture plated at similar cell densities (8×10^5 – 1.2×10^6 cells per dish). Transfections were carried out as described previously^{26,61} using 75 µg REF carrier DNA, 10 µg pEJ6.6 DNA and 1 µg of pSV2gpt DNA per 2×10^6 cells (2 dishes). After 24 h, the transfected cells were pooled. The REFs and Rat-1 cells were split in a ratio of 1:3 and 1:10, respectively. One day later half of the cultures were subjected to mycophenolic acid selection²⁷. Cultures were re-fed every 4 days. For plating in soft agar, 10⁶ cells were seeded 36 h after transfection into normal medium containing 0.3% low-melting agarose (Sea Plaque). Foci or colonies were counted 14–16 days after transfection. To test for *in vivo* growth potential, the transfected cells were collected 16–18 days after transfection, washed with phosphate-buffered saline (without Ca²⁺), and injected subcutaneously into 30–40-day-old nude mice. The mice had been irradiated (500 rad) 24 h before injection to eliminate natural killer cells. The cell content of one dish (5×10^6 cells) was used for a single injection. Cells which had been subjected to *Ecogpt* selection were adjusted to normal culture conditions 2–3 days before they were injected. This was done by re-feeding with *Ecogpt* selection medium²⁷ without mycophenolic acid and aminopterin. Finally, cells from one of these dishes were mixed with 5×10^6 untransfected 2° REFs and used for a single injection. The animals were observed for tumour formation on a weekly basis for 4 weeks. Tumours appeared between 7 and 10 days after injection.

from normal target cells for oncogenes *in vivo*. For example, the NIH 3T3 cells may, as a result of *in vitro* establishment, have acquired several alterations usually developed by a cell during its tumorigenic progression. These alterations may predispose the NIH 3T3 to tumorigenic conversion by a subsequent single-hit event. Thus, for the studies reported here we switched to rat embryo fibroblasts, reasoning that these cells more closely resembled normal targets of carcinogenic alteration.

Incomplete transformation by *ras* oncogene

The transforming gene used in our initial experiments was isolated from the human EJ bladder carcinoma cell line as a molecular clone termed pEJ6.6 (ref. 16). It represents a variant of the human c-Ha-ras1 proto-oncogene^{17–19}, and encodes a 21,000-molecular weight protein^{17,20,21}. This EJ *ras* oncogene stands as a model for other human oncogenes as it is a member of the *ras* gene family and is thus closely related to the Ki-ras and N-ras oncogenes found to be active in several different tumour types^{12,17–19,22–25}.

Copies of the cloned oncogene were introduced into recipient cells using the calcium phosphate transfection procedure of Graham and van der Eb²⁶. Mouse embryo fibroblasts were tested initially as recipients but were found to be unsatisfactory because of difficulties in detecting clones of stably transfected cells. Instead, we used secondary rat embryo fibroblasts (REFs) prepared from 12–14-day-old Fisher rat embryos. In the conditions of the focus assay (see Table 1 legend), no foci of morphologically altered cells were observed 14–21 days after transfection. These results were not due to an inability of the secondary (2°) REFs to take up and express exogenous DNA because when cultures of 10⁶ 2° REF cells were exposed to DNA of the *Ecogpt* clone²⁷ that serves as a dominant marker conferring resistance to growth inhibition by mycophenolic acid, 150 colonies were observable 14 days after transfection.

To examine further those cells in the culture that had taken up the EJ *ras* oncogene, but not yielded any obvious foci, we transfected the oncogene together with the *Ecogpt* marker and grew the cells in the presence of mycophenolic acid. (Cotransfected markers become incorporated together into competent cells in culture²⁸.) Of the resultant mycophenolic-acid-resistant colonies, 80–90% contained morphologically transformed cells (Fig. 1). Cells of the surrounding monolayer remained sparse and were unable to grow as they lacked resistance to the drug.

A second culture condition also permitted phenotypic expression of the introduced EJ *ras* oncogene. 2° REFs were exposed to DNA of the EJ *ras* oncogene, passaged, and introduced into soft agar suspension culture 36 h after transfection. The resultant 3° REF cultures formed 200 soft agar colonies per 10⁶ initially transfected cells. In contrast, no discernible colonies (>8 cells per colony) were found after agar culture of untransfected 3° REFs.

We compared the behaviour of these 3° REFs with that of cells of an established line that might behave similarly to NIH 3T3 cells. We chose the Rat-1 cell line that originated from Fisher rat embryo fibroblasts²⁹. When cultures of these cells were exposed to the EJ *ras* oncogene, large numbers of foci were seen (2,400 foci per 10⁶ cells) whether or not we used *Ecogpt* DNA-cotransfection followed by selection with mycophenolic acid (Table 1). A control experiment demonstrated that the transfected Rat-1 cells could form foci in conditions closely resembling those which did not permit focus formation by the 3° REFs. 500 EJ *ras*-transformed Rat-1 cells were mixed with 7.5×10^5 3° REFs and seeded in conditions identical to those of the earlier experiments; 20–30 foci were observed for every 100 transformed Rat-1 cells seeded into the REF monolayer. Thus, in virtually identical culture conditions, the transfected REFs were unable to form foci, while their Rat-1 counterparts did so quite efficiently (Table 2).

We found other contrasts between the behaviour of the transfected 3° REFs and Rat-1 cells. Oncogene-bearing cells could be recovered from *ras*-*Ecogpt*-cotransfected colonies of both types. The transformed Rat-1 cells yielded rapidly growing cell lines whereas colonies of the transformed 3° REFs entered cell crisis immediately on repassaging (they usually grew to a size of 500–5,000 cells, and then lost any further ability to divide).

Nude mice were inoculated with various transfected cultures and monitored for subsequent appearance of subcutaneous tumours. In one case the inoculated culture came from REF cells exposed to DNA of the EJ *ras* and *Ecogpt* clones and selected with mycophenolic acid: 80 resulting colonies, which contained in aggregate 5×10^4 transformed cells, were mixed with 5×10^6 untransfected 3° REFs before inoculation. This inoculum yielded only small, subcutaneous, cartilaginous nodules (average diameter 3 mm) 3 weeks after injection. In contrast, when 5×10^4 Rat-1 cells carrying the EJ *ras* oncogene were inoculated together with 2×10^6 untransfected Rat-1 cells,

the mixture induced rapidly growing fibrosarcomas which were easily observable after 1 week, and reached a size of 4 cm after 3 weeks.

It was clear that the EJ *ras* oncogene had only circumscribed powers. It was unable to impel transfected REFs into focal expansion in dense monolayer culture; the transfected cells had only limited proliferative ability; and they were not tumorigenic when seeded into host animals. However, activities that the oncogene lacked could be supplied by cellular functions existing in the Rat-1 cells prior to transfection. We concluded that the process of *in vitro* establishment provided the Rat-1 cells with functions that collaborate with the EJ *ras* oncogene to create a competent tumour cell.

Cooperative effects of *myc* and *ras*

We next wished to determine whether specific cellular functions, and by implication cellular genes, could work together with EJ *ras* to create a tumour cell. The cellular *myc* gene represented a candidate for such a cellular gene. As mentioned above, altered versions of this gene coexist with an active EJ *ras* oncogene in at least two different human tumour cell lines¹².

We tested the *myc* gene in the form of the oncogene carried by the avian MC29 virus. A molecular clone of its provirus, termed pv-*myc* (Fig. 2), was provided by J. M. Bishop³⁰. The high degree of evolutionary conservation of *myc*³¹⁻³³ suggested that the MC29 oncogene, while originating from the avian genome, might nevertheless be able to function in rat cells. A complete provirus clone was generated by joining a circularly permuted provirus clone²⁹ to an additional proviral segment at its right (3'-proviral) end. This created tandemly duplicated

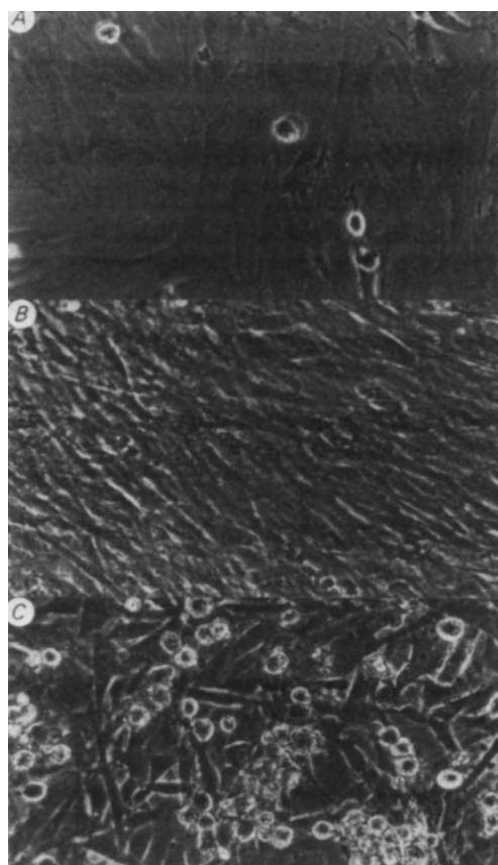


Fig. 1 Phase-contrast photomicrographs of REFs. ($\times 70$). **A**, Untransfected non-dividing 3° REFs forming a sparse monolayer in *Ecogpt* selection medium. **B**, Mycophenolic acid-resistant clone of REFs after transfection of pSV2gpt. **C**, Mycophenolic acid-resistant clone of REFs after cotransfection of pEJ6.6 and pSV2gpt, showing morphologically transformed cells. The DNA transfection was performed as described in Table 1 legend. Photographs were taken 14 days after transfection.

Table 2 Focus-forming ability of transfected Rat-1 cells or REFs after seeding into large excess of untransfected 3° REFs

Origin of cell clones seeded into REF monolayer			No. of foci per 100 seeded cells	
Original recipient cell	Introduced oncogene	Clone designation	Normal medium	<i>Ecogpt</i> selection medium
Rat-1	<i>ras</i>	1	20	9
Rat-1	<i>ras</i>	3	28	40
REF	<i>ras</i>	1 S	0	0
REF	<i>ras</i>	4 S	0	0
REF	<i>ras</i>	6 S	0	0
REF	<i>ras</i>	7 S	0	0
REF	<i>ras</i>	9 S	0	0
Rat-1	<i>ras-myc</i>	1	19	50
Rat-1	<i>ras-myc</i>	2	19	41
Rat-1	<i>ras-myc</i>	3	26	6
REF	<i>ras-myc</i>	1 N	38	38
REF	<i>ras-myc</i>	7 N	44	0
REF	<i>ras-myc</i>	3 S	55	68
REF	<i>ras-myc</i>	8 S	30	42

500 cells of a clone of Rat-1 cells or REFs which had been isolated after cotransfection of pSV2gpt and the human EJ c-Ha-*ras*1 clone, pEJ6.6 (designated *ras* clones), or after cotransfection of pSV2gpt, pEJ6.6 and pSVv-*myc* (designated *ras-myc*) (see also Tables 1 and 3) were mixed with 7.5×10^5 untransfected 3° REFs and seeded onto 10-cm dishes in either normal medium or *Ecogpt* selection medium. Focus formation was observed 6 days after plating. Foci were counted 12 days after seeding. N, Cell clone derived from a focus of transformed cells that had originally been isolated from a dish containing normal medium. S, Cell clone derived from a focus of transformed cells that had originally been isolated from a dish containing *Ecogpt* selection medium. Cells from all these clones were carried briefly in normal medium before being used in this assay (see also Table 1).

long terminal repeat (LTR) segments at the left (5'-proximal) and at the right (3'-proximal) ends of the provirus (termed psVv-*myc*; see Fig. 2). Initial experiments were designed to test the biological properties of these v-*myc* clones. Each DNA was applied to NIH 3T3 cells, to Rat-1 cells, or to 2° REFs, either alone or together with DNA of the *Ecogpt* clone. We observed no obvious effect on these transfected cultures.

Subsequent experiments tested the effects of the *myc* and EJ *ras* oncogenes introduced together into 2° REFs. In conditions in which either EJ *ras* or *myc* alone had no obvious effect on the monolayer cultures, the two genes together achieved a dramatic alteration of phenotype. Rapidly growing foci of morphologically altered cells (Fig. 3) became apparent within 8 days after transfection, whether or not the *Ecogpt*-cotransfection/mycophenolic acid selection protocol was followed. Such foci were able to expand into the surrounding monolayer, in all culture conditions.

When foci carrying the EJ *ras* and *myc* genes were picked, they yielded rapidly growing cultures (doubling time < 24 h) of morphologically transformed cells. This contrasted with the poor growth of cells carrying only the EJ *ras* gene (see above). These cotransfected cells were tumorigenic when introduced into nude mice or 12-day-old Fisher rats, yielding tumours that grew to a diameter of 1 cm 2 weeks after inoculation. Southern blot analysis of the DNAs from five of these lines confirmed the presence of multiple copies of both the transfected EJ *ras* and v-*myc* segments in these DNAs (data not shown).

These results led us to the tentative conclusion that the effects of *in vitro* establishment could be mimicked, at least in part, by introduction of the active *myc* oncogene. Both establishment and *myc* acquisition made the REFs highly reactive to the transforming effects of the EJ *ras* gene. However, a subtle and potentially significant difference emerged. The EJ *ras-myc* cotransfectants formed a tumour that reached a static size of 2 cm after 3 weeks of rapid expansion in the nude mouse hosts. In contrast, the established Rat-1 cells carrying the EJ *ras* gene induced tumours that expanded until they killed the host. We

interpret this as follows: the *myc* gene greatly enhances the phenotype created by the EJ *ras* oncogene, but is nevertheless unable to fully mimic the cellular traits conferred by establishment and immortalization *in vitro*.

Further characterization of *myc* gene

The above results caused us to examine in more detail those properties of the *myc* clone that allowed it to cooperate with the EJ *ras* oncogene. We showed that the LTRs, each one of which carried promoter and enhancer sequences³⁴, were not alone responsible for the observed activity of the clone, by deleting a 1.6-kilobase (kb) *SacI* fragment that lies in the protein-coding sequences of the *gag-myc* viral oncogene within pSV-v-*myc* (Fig. 2). When cotransfected with the EJ *ras* oncogene DNA, this modified provirus induced no detectable phenotype. Thus, activity of the *myc* clone depended on synthesis of part or all of the *gag-myc* protein.

Because the v-*myc* gene originated from the chicken genome, we thought that its activities might reflect the idiosyncracies of an avian gene acting in a foreign cellular environment. Therefore, we developed an active clone of a rodent *myc* gene. A mouse *myc* clone, isolated by Shen-Ong and colleagues³⁵, originated from DNA of the mouse plasmacytoma MOPC 315, in which a chromosomal translocation caused a *myc*-immunoglobulin gene juxtaposition. This immunoglobulin-*myc* hybrid

clone was modified by removing the immunoglobulin gene and placing the remaining *myc* segment under control of an early simian virus 40 (SV40) transcriptional promoter present in the SV2 *gpt* vector²⁷. This construction (termed pSVc-*myc*-1; Fig. 2) placed the transcriptional promoter and translation initiation codon 40 nucleotides apart.

This SV40-*myc* chimera was indistinguishable from the avian MC29 provirus in its ability to cooperate with a cotransfected EJ *ras* oncogene (Table 3). Optimal expression of this activity seemed to require that the transcription of the *myc* gene be driven directly by the SV40 promoter. When this promoter was introduced downstream of the gene, and in the opposite transcriptional polarity (pSVc-*myc*-2; Fig. 2), the number of colonies induced on cotransfection with a *ras* clone decreased by a factor of 10. A c-*myc* clone (termed pc-*myc*; Fig. 2) lacking any added promoter/enhancer showed no biological activity at all in the cotransfection assay (Table 3). Because of the functional equivalence of the avian and murine *myc* clones, we used the avian MC29-derived construct in further studies (see below).

Complementation groups

The ability of the EJ *ras* and *myc* genes to cooperate with one another led us to test whether other genes could be placed in complementation groups based on their activities in cooperating with either EJ *ras* or *myc* in the REF focus assay. An initial test measured the ability of the N-*ras* oncogene to behave like the Ha-*ras* oncogene (EJ *ras*) in its ability to cooperate with *myc*. A biologically active clone of this gene was isolated recently in this laboratory from HL 60 cells (J. M. Cunningham, in preparation). The two genes encode similar proteins. As expected, the N-*ras* gene behaved identically to the Ha-*ras* gene in its ability to cooperate with *myc* (Table 3). The two *ras* genes were therefore placed in the same complementation group.

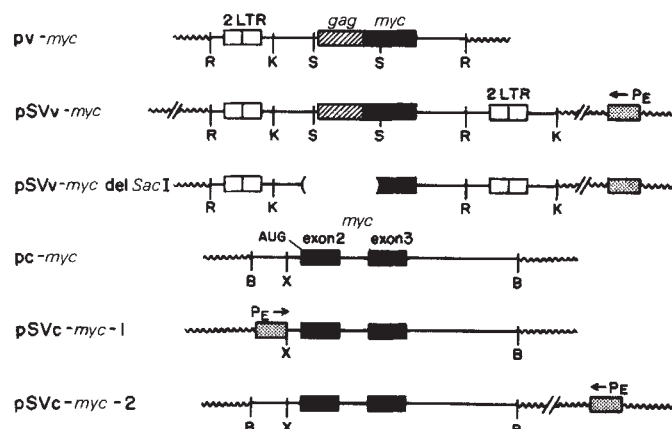


Fig. 2 Schematic diagram of plasmids containing various *myc* genes used in the cotransfection of 2° REFs: pv-*myc* contains a permuted copy of the avian MC29 provirus DNA (5.5 kb)³⁰ (given by J. M. Bishop). To provide a polyadenylation signal, the circularly permuted provirus clone was completed by duplication of the 5'-proximal 1.1-kb *EcoRI*-*KpnI* restriction fragment at the 3'-proximal end of the provirus. To achieve this, the *EcoRI*-*KpnI* fragment, which contains the two LTRs, was subcloned into the *EcoRI* and *KpnI* sites of pSV2gpt²⁷. In a second step, the *EcoRI* fragment containing the entire permuted proviral DNA was ligated into the modified vector. In the resulting plasmid, termed pSVv-*myc*, the SV40 early promoter (P_E) has opposite polarity to the direction of transcription of the proviral DNA. The pSVv-*myc* del plasmid, a further modification of the pSVv-*myc* described above, had a 1.6-kb *SacI* fragment deleted by cleavage with this enzyme and recircularization by ligation. This deletion removes a large portion of the sequences which encode the *gag-myc* fusion protein. pc-*myc* (given by M. Cole) contains the second and third exons of the cellular mouse *myc* gene within a 5.6-kb *Bam*HI fragment. It originated from DNA of the mouse plasmacytoma MOPC 315, in which this *myc* gene has been translocated into the immunoglobulin C_α locus³⁵. These immunoglobulin sequences were removed during construction of pc-*myc*. To express the mouse cellular *myc* gene of the pc-*myc* plasmid, we created a plasmid, termed pSVc-*myc*-1, in which transcription of the *myc* gene is driven by the SV40 early promoter (P_E). By linker ligation, an *Xba*I site was added to the *Hind*III site of pSV2gpt²⁷, then the 4.8-kb *Xba*I-*Bam*HI fragment of the pc-*myc* insert was cloned into this vector, separating the transcriptional start and AUG codon by 40 nucleotides. pSVc-*myc*-2 contains the *Bam*HI insert of pc-*myc* in pSV2gpt in opposite transcriptional orientation to the SV40 early promoter/enhancer sequences.

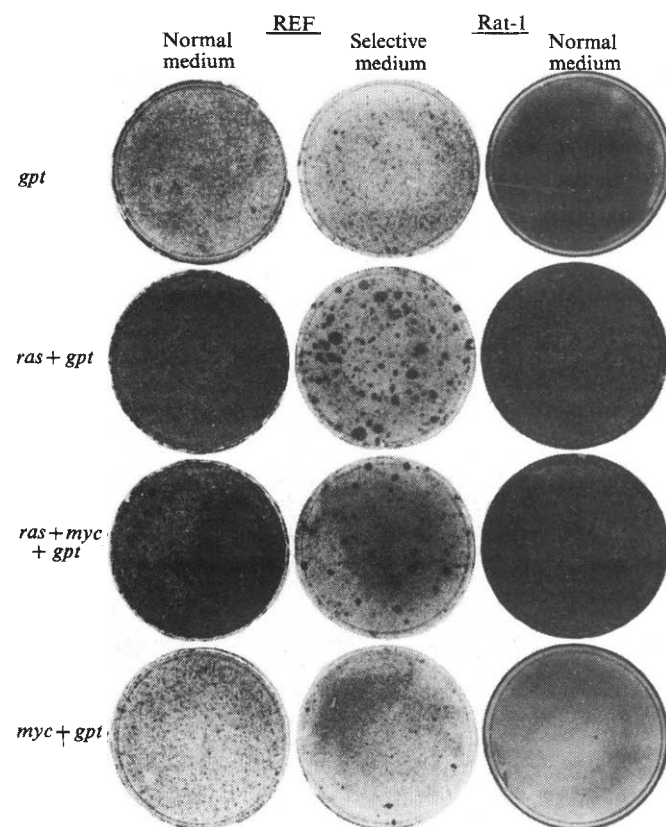


Fig. 3 Crystal violet-stained 3° REF and Rat-1 cell cultures photographed 14 days after transfection. Examples of the focus assays described in Tables 1 and 3 are shown. *gpt* Represents transfection of pSV2gpt DNA; *ras*, transfection of pEJ6.6 DNA; *myc*, transfection of pSVv-*myc* DNA.

Table 3 Complementation of transformation by cotransfection of different oncogenes into REFs

Transfected oncogene clones	Normal medium	Ecogpt Selection medium		Tumorigenicity in nude mice (no. of tumours/ no. of injections)
	Foci per 10 ⁶ cells	No. of colonies per 10 ⁶ cells	% Of colonies with morpho- logically transformed cells	
pSVv-myc	0	100	0	0/7
pSVc-myc-1	0	120	0	0/6
pEJ6.6+pSVv-myc	200	200	80	10/10
pEJ6.6+pSVv-myc del SacI	0	200	80	0/9
pEJ6.6+pv-myc	80	200	80	7/7
pEJ6.6+pc-myc	0	200	80	0/11
pEJ6.6+pSVc-myc 1	220	220	80	9/9
pEJ6.6+pSVc-myc 2	25	200	80	6/6
Φ N-ras*	0	140	15	0/5
Φ N-ras*+pSVv-myc	25	140	20	5/5
pPyMT1	0-10	75	15	0/6
pLT214	0	150	0	0/5
pPyMT1+pSVv-myc	60	100	15	2/5
pLT214+pSVv-myc	0	100	0	0/3
pEJ6.6+pPyMT1	0-10	70	10	0/5
pEJ6.6+pLT214	200	200	80	6/6

Cultures of 2° REFs were transfected by combinations of different cloned oncogenes as indicated. In each transfection, 1 µg of pSV2gpt per 2×10^6 cells was transfected together with 10 µg of each oncogene-carrying plasmid. Focus, colony, and tumorigenicity assays are described in detail in Table 1 legend. The data represent a mean of several experiments for each test.

* 2 µg of phage DNA were used.

We then tested the two viral genes encoding the middle- and large-T antigens of polyoma virus, isolated as separate molecular clones by Kamen and colleagues^{36,37}. As reported by Rassoulzadegan *et al.*⁶, these genes confer distinct and separable phenotypes on rat cells. The middle-T antigen induces morphological alteration and anchorage independence, while the large-T antigen affects serum dependency and cell immortalization.

Transfection of the middle T clone pPyMT1 either alone or together with the EJ *ras* DNA induced no obvious foci in our culture conditions. In contrast, cotransfection of middle-T and *myc* allowed outgrowth of dense foci. When these cultures were tested for tumorigenicity, they were found to induce tumours in only two of five animals. The remainder had only small nodules at the site of inoculation. Thus, middle-T behaved, in this experiment, similarly to but not identically with EJ *ras*, and can tentatively be assigned to the same complementation group as the *ras* oncogenes.

Manipulation of the large-T oncogene clone was more difficult. Initial experiments showed that it strongly inhibited establishment of cotransfected genes (data not shown), which may have been due to the presence of the polyoma replication origin together with the gene for the entire large-T antigen. This might allow the large-T antigen to trigger repeated rounds of viral DNA replication and in turn create a cytopathic effect. To avoid this problem, we chose a clone, termed pLT214, that carries a truncated version of the large-T antigen⁶. This clone, a gift from R. Kamen and colleagues, allowed synthesis of the N-proximal half of the large-T antigen implicated in alteration of serum dependency and immortalization⁶. The encoded large-T antigen lacked the C-proximal half, which is required for mediating viral DNA replication.

Transfection of the truncated large-T-antigen clone alone had no obvious effect on morphology of the 3° REF monolayer. The same result was obtained on cotransfection of this clone with the *myc* clone, pSVv-myc. However, when this altered large-T clone was cotransfected with EJ *ras*, dense foci appeared after 10 days. When cotransfected cultures were inoculated into six nude mice, rapidly growing tumours were found in all of them. These tumours reached a diameter of 4 cm after 3 weeks. Unlike the *myc*-EJ *ras*-induced tumours which stopped growing after reaching a diameter of 2 cm, these tumours continued to grow until they killed the host animals.

These data indicate that the truncated large-T gene behaves like *myc* in its ability to cooperate with a *ras* gene. However, the two genes do not function identically: *myc* allows growth of a large tumour of limited size, while the truncated large-T allows unlimited tumour growth.

Generality of experimental model

The present results raise several questions. Do these results apply only to fibroblast transformation induced by an arbitrarily chosen set of oncogenes? Or are the present results representative of processes that lead to many kinds of tumours?

One suggestion of possible generality comes from examination of the cellular oncogenes used here. *ras* and *myc* genes have been implicated as determinants in a wide variety of tumours. Thus active *ras* oncogenes have been found in bladder, colon, lung, pancreatic and skin carcinomas, neuroblastomas, sarcomas, and at least three types of haematopoietic malignancies^{17-19,23-25,38}. The cellular Ki-*ras* and Ha-*ras* genes encode almost identical gene products^{39,40}. Preliminary sequence analysis of a third member of this gene family, termed N-*ras*^{12,24,25}, indicates encoded amino acid sequences very similar to those of the other two members of the gene family (J. M. Cunningham, in preparation). We consider it likely that all members of this gene family act in a similar, if not identical fashion, and that the EJ c-Ha-*ras* 1 allele used in most of the present experiments represents a good model of all members of this gene family.

The *myc* gene is similarly involved in a range of tumours. It is associated with haematopoietic diseases such as myelocytomatosis⁴¹, myeloid leukaemia⁴²⁻⁴⁴, bursal lymphomas^{9,10}, Burkitt's lymphomas^{42,45,46} and plasmacytomas^{35,42,46-48}. Its involvement in non-haematopoietic disease is also well documented. Viruses that transduce the v-*myc* oncogene are known to induce kidney, pancreatic and liver carcinomas, and mesotheliomas⁴¹, and recent work shows the gene to be present in amplified copy number in a neuroendocrine tumour of the colon⁴⁹.

These diverse data suggest that neither type of gene is a 'tissue-specific' oncogene having a narrow range of tissue tropism. Instead, each type of gene appears to be competent in a variety of cellular environments. Moreover, the creation of disparate types of tumours seems to depend on common

molecular mechanisms frequently involving *ras* and/or *myc* oncogene activation. Because of the parallel behaviours of many tissues to these oncogenes, we believe that the fibroblasts used in this study represent credible models for a variety of target cells present in many tissues throughout the body.

Validity of the focus assay

The particular REF monolayer culture conditions used here do not allow focus induction by a *ras* oncogene (EJ *ras* or HL 60 N-*ras*) acting alone. However, by varying conditions, such as the frequency of passage and seeding density, we have been able to reveal focus-inducing ability by *ras* or polyoma middle-T oncogenes acting alone (unpublished results). Therefore, the conditions used in the present experiments have been chosen to demonstrate the limited powers of the singly acting oncogenes.

This choice of focus-assay conditions is vindicated by the use of other, independent measurements of oncogene activity. These other measurements—ability to grow in long-term culture and tumorigenicity—yield similar conclusions to those obtained from use of the focus assay. In all cases, the oncogenes of the *ras* group emerge as genes having potent but circumscribed abilities.

ras and *myc* act differently

The present results show that a *ras* oncogene is able to alter the morphology of REFs in monolayers in which the growth of surrounding cells has been suppressed, and can also induce colony growth in soft agar. However, the *ras* oncogene is unable to induce obvious foci in the midst of densely growing normal cells, and the REFs transformed by *ras* oncogene exhibit very limited proliferative and tumorigenic abilities. This would seem to contrast with the tumorigenic abilities of Ha- and Ki-murine sarcoma viruses, that transduce very closely related *ras* genes⁵⁰ and appear to be able to induce tumours with single-hit kinetics. We suggest that the *in vivo* steps in viral tumorigenesis are poorly understood, and may depend on additional cellular alterations beyond the acquisition of an exogenous oncogene.

The *myc* oncogene, when acting alone, has no effects on REFs that we can discern. Thus we cannot confirm other reports of rodent fibroblast transformation by the avian *myc* oncogene^{51,52}. We suggest that the reported transformations may stem from rare events that depend on secondary, cooperating cellular alterations.

One point that emerges from the present work is that *ras* and *myc* oncogenes act differently, because together they are able to achieve phenotypes that neither is able to achieve alone. It would seem that their modes of action differ qualitatively, and that each impinges on a distinct cellular target. The study of the *myc* oncogene has been hampered until now by the absence of any induced phenotype in mammalian cells. The cotransfection test described here may provide a useful technique for assaying the *myc* oncogene and related genetic segments.

Effects of *in vitro* establishment

It would appear that the processes of establishment and immortalization lead to development of cellular functions that are able to cooperate with the *ras* oncogene. We do not know how many separate cellular functions are required for development of the fully tumorigenic phenotype. However, it appears that the established cells carry all necessary functions except those provided by the *ras* oncogene, whose introduction produces full transformation with single hit kinetics.

Previously we had chosen to work with the established NIH 3T3 cells because of their efficient ability to take up and fix murine leukaemia virus DNAs¹⁴. Now it has become apparent that these cells also possess additional traits that make them unusually reactive to introduced oncogenes. Thus these cells are useful as indicators of the presence of certain trans-

forming oncogenes; they cannot be considered good models of normal target cells in the animal.

In experiments not described here we have introduced the (EJ) c-Ha-*ras*1 oncogene into Sprague-Dawley 2° REFs and only subsequently undertaken establishment of the transformed cells *in vitro*. These cells have acquired the tumorigenic phenotype. We conclude that the temporal order of the two events—*ras* oncogene acquisition and establishment—is not important for the final result.

Designation of separate functional classes

The *myc* oncogene introduction yields a phenotype which mimics in part the effects of *in vitro* establishment; in either case the addition of a *ras* oncogene creates the traits of focus formation and tumorigenesis. This does not mean that the introduction of *myc* and establishment create identical changes in cellular behaviour. Thus, we have no evidence that *myc* is able to immortalize cells, nor do we know whether the process of establishment/immortalization depends on activation of *myc*-like cellular genes. Nonetheless, the two factors are operationally very similar. This allows us to construct a complementation group in which we place *myc* and those functions subsumed under the term 'establishment'. The further inclusion of the large-T in this class might have been anticipated, even before the present results, since this gene had been implicated previously in establishment/immortalization functions⁶.

Recently, we have been informed of experiments indicating cooperation between a *ras* oncogene and the E1a oncogene of adenovirus 5 (ref. 62). Our own, preliminary experiments confirm that. Such a result is consistent with an earlier report showing that E1a acts, like polyoma large-T, to induce immortalization⁷. We assign E1a to the functional class that includes *myc*, large-T antigen, and establishment.

The other class of transforming oncogenes presently includes Ha-*ras*, N-*ras* and the polyoma middle-T. We note that this class, implicated in morphological alteration and anchorage independence, encodes gene products that are localized to the plasma membrane⁵³⁻⁵⁵. In contrast, the identified genes of the other class (*myc*, large-T and E1a) specify proteins that bind to nuclear structures⁵⁶⁻⁵⁹. This suggests that these functional groupings may reflect underlying common molecular mechanisms, including interaction with common cellular targets.

Finally, how does the activation of multiple oncogenes relate to the observed multi-step process of carcinogenesis? The activation of each oncogene appears to derive from a genetic alteration of low probability. Two of the steps normally required for tumour formation may represent the activation of a *ras*-like and a *myc*-like gene. It is not yet obvious whether these two steps will suffice to convert a normal cell into a fully competent tumour cell. Other steps may be required also. Thus, the cells carrying *myc* and *ras* oncogenes seed tumours that usually reach a large but static size. This contrasts with the behaviour of cells that have acquired a *ras* oncogene and have undergone establishment. Such cells seed tumours of apparently unlimited growth capacity. Perhaps the process of establishment/immortalization yields cellular functions beyond those achieved by *myc* alone. In that case, a third distinct gene may collaborate with *ras* and *myc* in creating the total tumorigenic phenotype. It would seem that the number of distinct steps in tumorigenesis is limited, and that each of these may soon be described at the molecular level.

We thank Bob Kamen for useful discussions and for his generosity in providing T-antigen clones; J. Michael Bishop and Michael Cole for their gifts of *myc* clones; Hartmut Beug for particularly useful suggestions; Jennifer Clark, Lalitha Vaidyanathan and Ann D. Rosen for technical assistance; and Susan Vazakas for her help in preparing the manuscript. H.L. is supported by the Deutsche Forschungsgemeinschaft. This work was supported by grants CA26717 and CA14051 of the US NCI.

Received 20 June; accepted 7 July 1983.

1. Cairns, J. *Nature* **255**, 197-200 (1975).
2. Vogt, M. & Dulbecco, R. *Proc. natn. Acad. Sci. U.S.A.* **49**, 171-179 (1963).
3. Berenblum, I. & Shubik, P. *Br. J. Cancer* **1**, 388 (1947).
4. Emmelot, P. & Scherer, E. *Cancer Res.* **37**, 1702-1708 (1977).
5. Crawford, B. D., Barrett, J. C. & Ts'o, P. O. P. *Molec. cell. Biol.* **3**, 931-945 (1983).
6. Rassoulzadegan, M. *et al. Nature* **300**, 713-718 (1982).
7. Houweling, A., van den Elsen, P. J. & van der Eb, A. J. *Virology* **105**, 537-550 (1980).
8. van den Elsen, P. J., de Pater, S., Houweling, A., van der Veer, J. & van der Eb, A. *Gene* **18**, 175-185 (1982).
9. Hayward, W. S., Neel, B. G. & Astrin, S. M. *Nature* **290**, 475-480 (1981).
10. Payne, G. S., Bishop, J. M. & Varmus, H. E. *Nature* **295**, 209-214 (1982).
11. Cooper, G. M. & Neiman, P. E. *Nature* **292**, 857-858 (1981).
12. Murray, M. J. *et al. Cell* (in the press).
13. Murray, M. J., Shilo, B.-Z., Shih, C. & Weinberg, R. A. *Cell* **25**, 355-361 (1981).
14. Smotkin, D., Gianni, A. M., Rozenblatt, S. & Weinberg, R. A. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4910-4913 (1975).
15. Aaronson, S. A. & Todaro, G. J. *J. Cell Biol.* **72**, 141-148 (1968).
16. Shih, C. & Weinberg, R. A. *Cell* **29**, 161-169 (1982).
17. Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
18. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-479 (1982).
19. Santos, E., Tronick, S., Aaronson, S. A., Pulciani, S. & Barbacid, M. *Nature* **298**, 343-347 (1982).
20. Shih, T. Y., Weeks, M. O., Young, H. A. & Scolnick, E. M. *Virology* **96**, 64-79 (1979).
21. Tabin, C. J. *et al. Nature* **300**, 143-149 (1982).
22. Perucho, M., Hanahan, D. & Wigler, M. *Cell* **22**, 309-317 (1980).
23. Pulciani, S. *et al. Nature* **300**, 539-542 (1982).
24. Hall, A., Marshall, C. J., Spurr, N. K. & Weiss, R. A. *Nature* **303**, 396-400 (1983).
25. Shimizu, K., Goldfarb, M., Peruchio, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 383-387 (1983).
26. Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
27. Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072-2076 (1981).
28. Perucho, M., Hanahan, D. & Wigler, M. *Cell* **22**, 309-317 (1980).
29. Freeman, A. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 2415-2419 (1973).
30. Vennstrom, B., Moscovici, C., Goodman, H. M. & Bishop, J. M. *J. Virol.* **39**, 625-631 (1981).
31. Watson, D. K., Reddy, E. P., Duesberg, P. H. & Papas, T. S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2146-2150 (1983).
32. Alitalo, K. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 100-104 (1983).
33. Colby, W. W., Chen, E. Y., Smith, D. H. & Levinson, A. D. *Nature* **301**, 722-725 (1983).
34. Laimins, L. A., Khoury, G., Gorman, C., Howard, B. & Gruss, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6453-6457 (1982).
35. Shen-Ong, G. L. C., Keath, E. J., Piccoli, S. P. & Cole, M. D. *Cell* **31**, 443-452 (1982).
36. Treisman, R., Novak, U., Favalaro, J. & Kamen, R. *Nature* **292**, 595-600 (1981).
37. Tyndall, C., LaMantia, G., Thacker, C. M., Favalaro, J. & Kamen, R. *Nucleic Acids Res.* **9**, 6231-6250 (1981).
38. Taparowsky, E. *et al. Nature* **300**, 762-765 (1982).
39. Tsuchida, N., Ryder, T. & Ohtsubo, E. *Science* **217**, 937-939 (1982).
40. Dhar, R. *et al. Science* **217**, 934-937 (1982).
41. Beard, J. W. in *Viral Oncology* (ed. Klein, G.) 55-88 (Raven, New York, 1980).
42. Adams, J. M., Gerondakis, S., Webb, E., Corcoran, L. & Cory, S. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
43. Collins, S. J. & Groudine, M. *Nature* **298**, 679-681 (1982).
44. Dalla Favera, R., Wong-Staal, F. & Gallo, R. C. *Nature* **299**, 61-63 (1982).
45. Dalla Favera, R., *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7824-7827 (1982).
46. Taub, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 7837-7841 (1982).
47. Crews, S., Barth, J., Hood, L., Prehn, J. & Calame, K. *Science* **218**, 1319-1321 (1982).
48. Marcu, R. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 519-523 (1983).
49. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1707-1711 (1983).
50. Gross, L. *Oncogenic Viruses*, 931 (Pergamon, Oxford, 1970).
51. Quade, K. *Virology* **98**, 461-465 (1979).
52. Copeland, N. G. & Cooper, G. M. *J. Virol.* **33**, 1199-1202 (1980).
53. Willingham, M. C., Pastan, I., Shih, T. Y. & Scolnick, E. M. *Cell* **19**, 1005-1014 (1981).
54. Ito, Y. *Virology* **98**, 261-266 (1979).
55. Schaffhausen, B. S., Dorai, J., Arakere, G. & Benjamin, T. L. *Molec. cell. Biol.* **2**, 1187-1198 (1982).
56. Donner, P., Greiser-Wilke, I. & Moelling, K. *Nature* **296**, 262-266 (1982).
57. Abrams, H. D., Rohrschneider, L. R. & Eisenman, R. N. *Cell* **29**, 427-439 (1982).
58. Ito, Y., Spurr, N. & Dulbecco, R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1259-1263 (1977).
59. Feldman, L. T. & Nevins, J. R. *Molec. cell. Biol.* **3**, 829-838 (1983).
60. Pollack, R., Risser, R., Coulon, S. & Rifkin, D. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4792-4796 (1974).
61. Andersson, P., Goldfarb, M. P. & Weinberg, R. A. *Cell* **16**, 63-75 (1979).
62. Ruley, H. E. *Nature* **304**, 602-606 (1983).

Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture

H. Earl Ruley

Cold Spring Harbor Laboratory, Cold Spring Harbor, PO Box 100, New York 11724, USA

The polyoma virus middle-T and the T24 Harvey ras1 genes are individually unable to transform primary baby rat kidney cells. Adenovirus early region 1A provides functions required by these genes to transform primary cells following DNA-mediated gene transfer. These results suggest that separate establishment and transforming functions are required for oncogenic transformation of primary cells in culture.

THE oncogenic DNA viruses encode proteins that are able to transform cells grown in tissue culture¹. Virus-transformed cells, which can be isolated by their ability to form dense foci on monolayers of untransformed cells, typically retain and express all or some of the viral genes that are expressed early during lytic infection. Both primary and established cells can be transformed by polyoma virus (Py) and human adenovirus (Ad). However, the viral functions required to transform these cell types are different, at least in some cases.

Transformation of primary cells requires at least two separate functions. The first, an establishment function, is concerned with immortalization of cells, while the second, the transformation function, is required for full expression of an oncogenic phenotype²⁻⁵. Thus, establishment functions expressed by adenovirus early region 1A (E1A) or portions of the polyoma large-T antigen lead to the ability of primary cells to grow indefinitely in culture^{2,3}. Additional functions expressed by adenovirus early region 1B (E1B) or the polyoma virus middle-T antigen result in phenotypic changes characteristic of oncogenic transformation²⁻⁵ such as anchorage-independent cell growth and the ability to form tumours when cells are

transplanted into syngeneic animals.

By contrast, transformation of established cell lines can require fewer viral functions. Thus, expression of the polyoma middle-T antigen alone is sufficient to transform a variety of cell lines⁵. Apparently, such cell lines constitutively express establishment functions that can substitute for those of the virus. The interaction between establishment and transforming functions is poorly understood, as is the mechanism by which they combine to elicit the transformed phenotype.

Recently, several genes have been isolated from cell lines established from human tumours, that have the ability to cause morphological transformation of the mouse NIH 3T3 cell line⁶⁻¹⁰. However, the ability of cellular oncogenes to transform cultured primary cells has not been critically addressed. Given the requirement for at least two viral functions for transformation of primary cells by polyoma and adenoviruses, it seemed quite likely that cellular oncogenes might also require additional functions. A failure to transform primary cells could be explained if oncogenes that score in the NIH 3T3 assay carry transforming functions but lack establishment functions.

Table 1 Transformation of primary BRK cells with cloned viral and cellular genes

Plasmid	Experiment		
	I	II	III
None (carrier alone)	0	0	0
Ad-2 E1A	(8)	(17)	(19)
T24 Ha-ras1	0	0	0
PyMT	0	0	0
Ad-2 E1B	0	0	0
Ad-2 E1A + T24 Ha-ras1	23	57	130
Ad-2 E1A + PyMT		10	65
Ad-2 E1A + Ad-2 E1B		24	22
P3 Ha-ras1	0	0	0
Ad-2 E1A + P3 Ha-ras1	0	0	0
pLA8	121	167	
PyBamL		48	45
Ad-2 E1A + PyBamL		41	67
PydXba		14	32
Ad-2 E1A + PydXba		26	40

Plasmid DNAs were transfected into primary cells by co-precipitation with calcium phosphate as described by Wigler *et al.*²². Aliquots (0.5 ml) of DNA-calcium phosphate precipitate containing 250 ng of each plasmid and 10 g rat carrier DNA was added to 60-mm dishes of primary BRK cells. Cells were fixed with methanol/acetone (1:1) and stained with Giemsa 3–8 weeks post-transfection. The total number of foci on four plates is indicated in each experiment. Cells transfected with Ad-2 E1A plasmids did not form dense foci (see Fig. 1) but formed colonies of cells having extended growth potential compared with their non-transformed counterparts. The number of colonies of this type are listed in parentheses except when Ad-2 E1A was co-transfected with other plasmids, in which case they were not included in the totals. Plasmids contained the following viral and cellular sequences: pHEB4 (Ad-2 E1A, nucleotides 269–1,571), pXX6 (Ad-2 E1B, 3.7–15.8 map units), pLA8 (the left-end 9.1% of the Ad-2 genome, a transforming segment containing E1A and most of E1B); PyBamL (an intact Py genome); PydXba (a replication-defective Py genome lacking coding sequences for the carboxy-terminal domain of the large-T antigen, nucleotides 2,477–2,522); PyMT (a modified Py genome which expresses only the middle-T antigen protein)⁵; T24 Ha-ras1 (the activated Ha-ras1 gene isolated from T24 human bladder carcinoma cells)⁶ and P3 Ha-ras1 (a non-activated allele of Ha-ras1 isolated from a human placental DNA library)¹⁹. Py and Ad-2 nucleotide coordinates are from Soeda *et al.*²³ and Gingeras *et al.*²⁴, respectively. Plasmids were constructed or were provided by the following persons: pLA8 and pXX6, N. Stow; pHEB4, D. Solnick; PyBamL, M. Fried; P3 and T24, M. Wigler; PyMT, R. Kamen. Primary rat kidney cells were prepared by collagenase/dispase (Boehringer) treatment of kidneys from 6-day-old Fisher rats. $3\text{--}5 \times 10^5$ cells per 60-mm dish were plated 2 days before DNA transfections. Cultures were ~70% confluent at the time of transfection. Cultures were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5% fetal calf serum and antibiotics.

The present study investigated the ability of cloned viral and cellular genes to transform primary cultures of rat kidney cells. These genes were tested individually and in various combinations to determine whether viral transforming and establishment functions could be interchanged and whether viral establishment functions were required by cellular oncogenes to transform primary cells in culture.

Transforming activities

The ability of viral genes and cellular oncogenes to transform primary baby rat kidney cells (BRK) was tested following transfection with DNA-calcium phosphate co-precipitates. Transformation was monitored by the appearance of foci of rapidly growing cells 2–8 weeks post-transfection. Plasmids containing the following DNA sequences were tested singly and in combination: Ad-2 early regions 1A or 1B; the left-most 9% of Ad-2 DNA, a complete Py genome or a modified Py genome separately encoding the Py middle-T antigen species; the cellular oncogene (T24 Ha-ras1) isolated from T24 human bladder carcinoma cells (a gene related to the oncogene transduced by the Harvey murine sarcoma virus); and a normal form of the

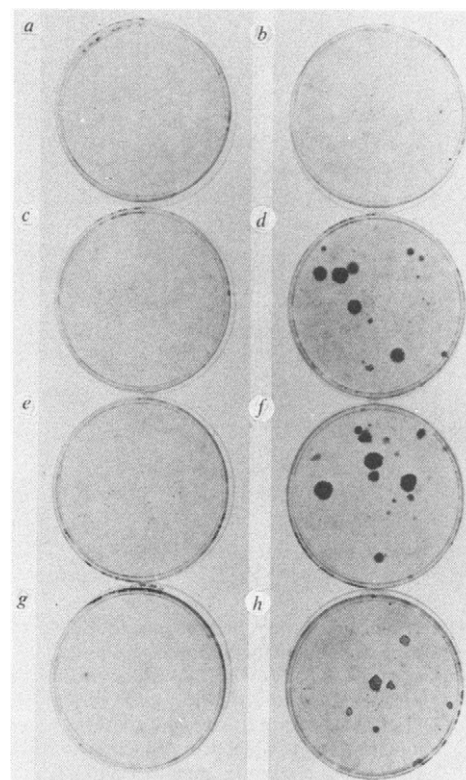


Fig. 1 Focus assay following transfection of BRK cells by transforming genes alone or with Ad-2 E1A. Cells were transfected with plasmid DNAs and fixed and stained at various times thereafter as described in Table 1 legend. *a*, No plasmid; *b*, Ad-2 E1A; *c*, T24 Ha-ras1; *d*, Ad-2 E1A + T24 Ha-ras1; *e*, PyMT; *f*, Ad-2 E1A + PyMT; *g*, Ad-2 E1B; *h*, Ad-2 E1A + Ad-2 E1B. The duration of each assay was 28 days except for *d*, which was 18 days.

same gene (P3 Ha-ras1), isolated from a library of human placental DNA.

Table 1 gives the results of several experiments. Control plasmids separately encoding the left-most 9% of the Ad-2 genome or a complete Py genome transformed BRK cells with similar efficiencies. By contrast, no transformed foci were detected following transfection of BRK cells with plasmids separately encoding the Py middle-T antigen species or Ad-2 E1B. These results are in agreement with those previously published from other laboratories^{4,11}. Table 1 shows that the T24 Ha-ras1 gene alone was also unable to transform primary rat kidney cells. This result is interesting in view of the fact that the T24 Ha-ras1 gene transforms both NIH 3T3 and Rat-1 established cell lines^{6,12}.

When plasmids carrying sequences coding for either the Py middle-T antigen or the T24 Ha-ras1 genes were co-transfected with plasmids containing Ad-2 E1A, foci were generated with frequencies similar to those obtained following co-transfection of BRK cells with Ad-2 E1A and E1B, the entire Py genome or the left-most 9% of the Ad-2 genome. However, P3 Ha-ras1 (a normal allele of the T24 Ha-ras1 gene) was nontransforming regardless of whether it was introduced into BRK cells together with Ad-2 E1A.

A focus assay following transfection of BRK cells with transforming genes alone or in combination with E1A is shown in Fig. 1. Transformed foci appeared against a background of normal fibroblasts that overgrew and displaced the monolayer of epithelial cells. While these primary cultures constitute a mixture of cell types, cotransfer of E1A is required for transformation regardless of the target cell type. The absolute requirement for E1A suggests that this region of the Ad-2 genome expresses an activity required for transformation of primary BRK cells by both viral and cellular transforming genes.

These results do not exclude the possibility that the T24 Ha-ras1 Py middle-T antigen genes transform BRK cells at an

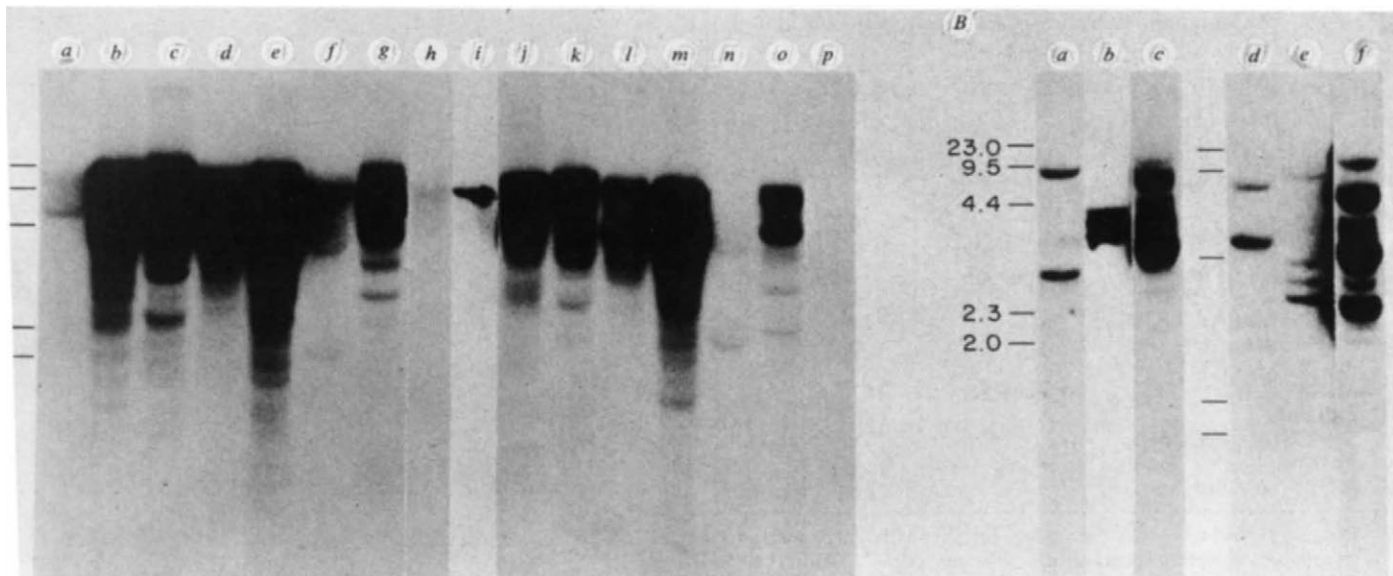


Fig. 2 Detection of sequences homologous to transforming plasmids in transformed cell DNA. High molecular weight cellular DNAs were cleaved with *Bam*HI and analysed by Southern blot hybridization²⁵. **A**, Filters were probed with ³²P-labelled plasmids²⁶ containing either T24 Ha-ras1 (lanes a–h) or Ad-2 E1A sequences (lanes i–p). Ad-5-transformed human embryonic kidney cell line, 293¹⁴ (a, i); Ad-2 + T24 Ha-ras1 transformed cell lines: Q1 (b, j), Q2 (c, k), Q3 (d, l), Q4 (e, m), D2 (f, n), S1 (g, o) and primary BRK cells (h, p). The sizes (in kilobase pairs) of some of the fragments generated by *Hind*III cleavage of gels DNA are indicated. The autoradiograms were exposed for 20 h except in lanes i and p, which were exposed for 3 and 5 days, respectively. **B**, Filters were probed with ³²P-labelled plasmids containing either Ad-2 E1A (lanes a–c) or Py middle-T antigen (lanes d–f) sequences. Ad-2 E1A + PyMT transformed cell lines: 19-5 (a, d), 1AMT-1 (b, e) and 1AMT-2 (c, f). Autoradiographic exposure was 18 h (c, f) and 72 h (a, b, d, e).

extremely low frequency which Ad-2 E1A enhances to detectable levels. However, if such is the case, E1A must enhance transformation by over a thousandfold to account for the cumulative difference obtained in multiple experiments (data not shown).

The results shown in Table 1 can be explained in several ways. Plasmids carrying E1A may enhance integration or expression of the co-transfected sequences; alternatively, functions expressed by E1A in addition to those of the transforming genes are required to transform primary BRK cells. While difficult to eliminate entirely, the following data argue strongly against the first two possibilities. First, the transforming efficiency of plasmids containing wild-type (PyBamL) or replication-defective (PyDXba) polyoma viral genomes is unaffected by co-transformation with E1A (Table 1). Moreover, co-transfection of E1A does not increase the amount of integrated plasmid DNA as judged by Southern blot analysis of DNA extracted from cells transformed by either plasmid (data not shown). Second, as shown below, there was no correlation between the levels of p21 (the proteins encoded by the *ras* genes) and RNA homologous to E1A in the transformed cells. This result does not exclude the possibility that the promoter of the T24 Ha-ras1 gene is non-functional in primary BRK cells and that an activity associated with E1A, acting either in *cis* or in *trans*, is required for T24 Ha-ras1 expression. However, this seems unlikely as the Py middle-T antigen gene is similarly dependent on E1A for transformation even though the Py early promoter is functional in BRK cells.

Transfected DNA expression

Transformed foci were picked using cloning cylinders and the clones expanded to mass culture. High molecular weight DNA was extracted and analysed by Southern hybridization for the presence of sequences homologous to the transfecting plasmids (Fig. 2). Considerable variation existed in the amount of homologous sequences present in different clones. While many transformants contained up to several hundred copies, several lines (S1, 19-5, 1AMT-1) retained only one or a few copies of the transfected DNA. These results are unremarkable as vari-

ation in the levels of acquired DNA sequences is a typical feature of cells that have been transfected by DNA-calcium phosphate co-precipitates¹³.

To determine whether E1A DNA sequences are expressed in E1A/T24 co-transformants, poly(A)-containing RNA, selected from six independently derived cell lines, was examined by Northern hybridization (Fig. 3). For comparison RNA was also isolated from an Ad-5 transformed cell line (293) containing a single insert of viral DNA¹⁴. All six cell lines contained different levels of RNA homologous to the E1A probe and in all cases the RNA species were larger than those expressed in 293 cells. This result was expected, as the co-transfected E1A plasmid (pHEB4) extends only to the viral *Hpa*I site (at nucleotide 1,571), and the polyadenylation signal for the E1A region is therefore deleted¹⁵. Thus, the larger RNAs probably represent hybrid transcripts that extend into flanking DNA sequences (either cellular or plasmid). Similar transcripts have been detected in cells transfected by Ad-5 *Hpa*IE fragment DNA².

The T24 Ha-ras1 gene encodes a 21,000-molecular weight protein referred to as p21. Similar proteins, expressed by other members of the *ras* gene family, have been found in all vertebrate cells examined thus far¹⁶. Transformants containing both T24 and E1A DNA plasmid sequences were labelled with ³⁵S-methionine and proteins related to p21 were immunoprecipitated and analysed by SDS-polyacrylamide gel electrophoresis. A broadly cross-reactive anti-p21 monoclonal antibody was used that precipitates the p21 products encoded by normal cellular *ras* genes¹⁷ as well as that encoded by the T24 Ha-ras1 gene¹⁸. The cellular Ha-ras and T24 Ha-ras1 gene products differ in that the glycine residue at position 12 of Ha-ras is replaced by valine in the T24 Ha-ras1 gene. This change, which is responsible for the transforming properties of the T24 Ha-ras1 gene^{18–20}, also results in a protein having lower electrophoretic mobility. The T24 Ha-ras1 gene product expressed in transformed cells can therefore be easily distinguished from the forms of p21 encoded by the endogenous rat genome¹⁸ (Fig. 4).

Two species of p21, thought to reflect processed and non-processed forms of the p21 proteins, were detected²¹. While all the E1A/T24 co-transformants express p21 species corres-

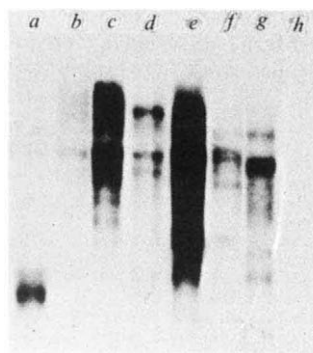


Fig. 3 Expression of Ad-2 E1A RNA in Ad-2 E1A+T24 Ha-ras1 transformants. Poly(A)-containing RNA was fractionated on 1.4% formaldehyde agarose gels²⁷, blotted to nitrocellulose²⁸ and probed with ³²P-labelled plasmid DNA containing Ad-2 E1A sequences (pHEB4, specific activity, 2×10^8 c.p.m. μg^{-1}). Ad-5 transformed human embryonic kidney cell line, 293 (a), Ad-2 E1A+T24 Ha-ras1 transformed cell lines: Q1 (b) Q2 (c), Q3 (d), Q4 (e), D2 (f), S1 (g) and Harvey murine sarcoma virus (Ha-MuSV)-transformed NIH 3T3 cells²⁹ (f). The autoradiogram for a, b and c–g was exposed for 120 and 18 h, respectively. Total cellular RNA was isolated by guanidium thiocyanate–hot phenol extraction³⁰. Poly(A)-containing RNA was enriched by one cycle of oligo(dT)–cellulose chromatography³¹.

ponding to the T24 Ha-ras1 gene product, the levels varied considerably among the different cell lines. In particular, lines Q1 and Q4 expressed high levels of the activated p21, whereas Q3 and D2 expressed levels comparable to the endogenous rat p21. p21 levels were not correlated with levels of E1A expression as estimated by Northern blot analysis (Fig. 3) and were only loosely correlated with gene copy number (Fig. 2). Nevertheless, all the cell lines were fully transformed in that they grew to high saturation densities and were tumorigenic in syngeneic rats (data not shown). Tumorigenicity was monitored after subcutaneous injection of 10^6 cells into each of four 14-day-old Fisher rats. Visible tumours were apparent in at least three of the animals within 3–4 weeks. Tumours grew progressively requiring killing of the animals after 5–8 weeks. These results indicate that (1) modest levels of the activated p21 are sufficient for transformation of primary cells following co-transfer of E1A, and (2) E1A does not function by activating expression of the T24 Ha-ras1 or cellular *ras* genes to high levels.

Discussion

Cloned viral and cellular genes were introduced into primary cultures of BRK cells to examine potential interactions leading to oncogenic transformation. The Py middle-T antigen, Ad-2 E1B and T24 Ha-ras1 genes transform primary BRK cells following co-transfer with Ad-2 E1A and fail to transform in the absence of E1A. The activity of E1A required by transforming genes to transform primary cells is unknown, but E1A does not seem to enhance either the integration or expression of co-transfected sequences.

It has been previously demonstrated that expression of E1A alone is able to immortalize primary cells². It seems likely that the E1A activity which complements transforming functions of the T24 Ha-ras1 and middle-T antigen genes is associated with functions which confer on primary cells the ability to grow indefinitely in culture. The link between establishment and transforming functions is further demonstrated by the fact that both the Py middle-T antigen and T24 Ha-ras1 genes readily transform a variety of established cell lines^{5,6,11,12}. This raises

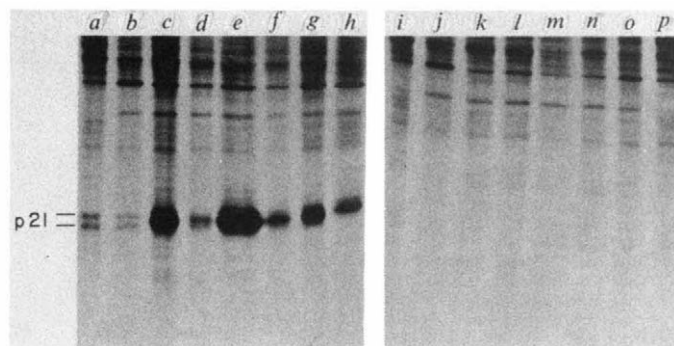


Fig. 4 Expression of p21 in Ad-2 E1A+T24 Ha-ras1 transformants. ³⁵S-methionine-labelled proteins were immunoprecipitated using anti-p21 monoclonal antibody (lanes a–h) or normal rat serum (lanes i–p). Ad-5-transformed human embryonic kidney cell line, 293 (a, i), primary BRK cells (b, j), Ad-2 E1A+T24 Ha-ras1-transformed cell lines: Q1 (c, k), Q3 (d, l), Q4 (e, m), D2 (f, n), S1 (g, o) and Ha-MuSV-transformed NIH 3T3 cells²¹ (h, p). Monolayer cultures grown in 6-cm dishes were washed in methionine-free DMEM and labelled for 4 h in 1 ml of methionine-free DMEM containing 500 μCi ³⁵S-methionine (NEN; 1,000 Ci mmol^{-1}); cells were washed with 50 mM Tris, pH 7.5, 0.15 M NaCl and lysed at 0°C, with 1 ml of buffer containing 1% Nonidet P-40 (NP40), 150 mM NaCl and 50 mM Tris pH 8.0. After 30 min the lysate was vortexed briefly and centrifuged at 100,000g for 30 min. Supernatant fractions containing 10^6 acid-precipitable c.p.m. were incubated with normal rat serum or a broadly cross-reactive anti-p21 rat monoclonal antibody (Y13–259)²⁵. Immune complexes were bound to rabbit anti-rat IgG (Cappel) previously adsorbed to 3.5 mg protein A–Sephacrose (Pharmacia) and recovered by centrifugation. After three washes (0.15 M NaCl, 0.05 M Tris pH 7.5, 5 mM EDTA, 0.05% NP40, 0.25% gelatin and 0.02% sodium azide), immune complexes were disrupted and electrophoresed in 15% polyacrylamide gels¹³. Gels were processed for fluorography and dried before autoradiography³².

the possibility that cellular establishment genes are activated in primary cells selected to grow continuously in culture.

The T24 Ha-ras1 gene was isolated from the T24 line of human bladder carcinoma cells by virtue of its ability to transform NIH 3T3 cells^{6–8}. However, the present study shows that the T24 Ha-ras1 gene is unable to transform primary BRK cells. Moreover, functions supplied by Ad-2 E1A can work in concert with the T24 Ha-ras1 gene to transform primary BRK cells following DNA-mediated gene transfer. P3 Ha-ras1 (a non-activated allele of the T24 Ha-ras1 gene) fails to transform with or without E1A. These results demonstrate that activation of Ha-ras1 is necessary but not sufficient for oncogenic transformation of BRK cells in culture. It will be of interest to determine whether activities analogous to those expressed by E1A are expressed in human tumour cells containing activated *ras* transforming genes.

The results presented here suggest that some viral and cellular transforming genes can be classified by function into two groups which define different steps required for the transformation of primary cells *in vitro*. Each step can be assayed independently by inducing different alterations in the ability of cells to proliferate in culture. The establishment step, by leading to the ability of primary cells to grow continuously in culture, appears to enable primary cells to respond to both serum growth factors and the action of transforming proteins. Recently, Land *et al.*³⁴ have shown that *c-myc* and polyoma virus large-T antigen can provide functions provided by EJ Ha-ras gene to transform primary rat embryo fibroblasts. These results, together with those presented here, suggest that polyoma and adenovirus early region proteins express functions similar to those of several oncogenes activated in human tumour cells.

I thank John Moomaw and Anne Devine for technical assistance, Mary Merle, Lisa Davis and Deborah Lukralle for help in preparing primary cell cultures, Ed Harlow for advice on immunoprecipitations, Terri Grodzicker for helpful discussions,

Received 26 July; accepted 3 August 1983.

1. Tooze, J. (ed.) *Molecular Biology of Tumor Viruses Pt 2* (revised) (Cold Spring Harbor Laboratory, New York, 1981).
2. Houweling, A., Van den Elsen, P. & Van der Eb, A. *Virology* **105**, 537–550 (1980).
3. Rassoulzadegan, M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
4. Van der Eb, A. J. *et al. Cold Spring Harb. Symp. quant. Biol.* **44**, 383–399 (1980).
5. Triesman, R., Novak, U., Favoloro, J. & Kamen, R. *Nature* **292**, 595–600 (1982).
6. Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404–409 (1981).
7. Shih, C. & Weinberg, R. A. *Cell* **29**, 161–169 (1982).
8. Pulciani, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2845–2849 (1982).
9. Shimizu, K., Goldfarb, M., Perucho, M. & Wigler, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 383–387 (1983).
10. Goubin, G., Goldman, D. S., Luce, J., Neiman, P. E. & Cooper, G. M. *Nature* **302**, 114–119 (1983).
11. Rassoulzadegan, M. *et al. Nature* **300**, 713–718 (1982).
12. Capan, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33–37 (1983).
13. Wigler, M. *et al. Cell* **16**, 777–785 (1980).
14. Aiello, L., Guilfoyle, R., Huebner, K. & Weinman, R. *Virology* **94**, 460–469 (1979).
15. Pericaudet, M., Akusjarvi, G., Virtanen, A. & Pettersson, U. *Nature* **281**, 694–696 (1979).

Joe Sambrook and Terri Grodzicker for critical readings of this manuscript, and Marilyn Goodwin for help in preparing this manuscript. This work was supported by NCI grant CA13106.

16. Langbeheim, H., Shih, T. Y. & Scolnick, E. M. *Virology* **106**, 292–300 (1980).
17. Furth, M. E., Davis, L. J., Fleurdelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294–304 (1982).
18. Tabin, C. J. *et al. Nature* **300**, 143–149 (1982).
19. Taparowsky, E. *et al. Nature* **300**, 762–765 (1982).
20. Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
21. Shih, T. Y. *et al. J. Virol.* **42**, 253–261 (1982).
22. Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373–1376 (1979).
23. Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. & Griffin, B. E. *Nature* **283**, 445–453 (1980).
24. Gingeras, T. *et al. J. biol. Chem.* **257**, 13475–13491 (1982).
25. Southern, E. J. *molec. Biol.* **98**, 503–517 (1975).
26. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237–251 (1977).
27. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
28. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).
29. Scolnick, E. M. & Parks, W. P. *J. Virol.* **13**, 1211–1219 (1974).
30. Feramisco, J. R., Smart, J. E., Burridge, K. & Thomas, G. P. *J. biol. Chem.* **257**, 11024–11031 (1982).
31. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
32. Lammli, U. K. *Nature* **227**, 680–685 (1970).
33. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83–88 (1974).
34. Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596–602 (1983).

LETTERS TO NATURE

Upper limit for mediation of baryon decay by slow magnetic monopoles

N. A. Porter, J. Clear, D. J. Fegan & G. C. MacNeill

Physics Department, University College, Belfield, Stillorgan Road, Dublin 4, Ireland

K. Gibbs & T. C. Weekes

Harvard-Smithsonian Center for Astrophysics, Whipple Observatory, Arizona 85640, USA

Rubakov¹ and Callan² have pointed out that grand unified theory magnetic monopoles will mediate the decay of baryons, and find that the cross-section for the process should be in the strong interaction region. Others^{3,4} find a smaller cross-section but agree that the process will occur. Severe constraints on the density of these monopoles in the Galaxy can be deduced from astrophysical arguments^{5–7}, but it is possible that local clustering due to the Sun⁸ may produce a greater density at the Earth, particularly for particles with $\beta < 10^{-4}$. Using an atmospheric Cerenkov detector, we describe here how to set a limit to the flux of monopoles on the Earth's surface at these low velocities.

An upper limit has already been set by a magnetic flux experiment^{9,10}, in which one candidate event has been seen. This limit is at present 0.1 monopoles $\text{m}^{-2} \text{day}^{-1} \text{sr}^{-1}$, (irrespective of velocity), which is several orders of magnitude greater than the astrophysical values. Experiments based on timing of ionizing particles have set values for $\beta > 10^{-4}$ (refs 11–13, and T. Takahashi, personal communication) and one experiment¹⁴ based on the Rubakov–Callan effect has given an upper limit flux of 7.5×10^{-3} monopoles $\text{m}^{-2} \text{day}^{-1} \text{sr}^{-1}$ for values of β lying between 5×10^{-4} and 5×10^{-2} . There seems to be no experimental limit for $\beta < 10^{-4}$, where clustering in the Solar System may be strong.

We present here measurements made with an atmospheric Cerenkov detector, normally used for γ -ray astronomy. While the system is probably more susceptible to spurious signals than are underground detectors built for baryon-decay experiments, it operates in a totally different time domain because the medium is air rather than water or iron. It uses a substantial mass of air, approaching 1,000 tonnes, and has a high area-to-mass ratio. A positive effect could not be claimed with confidence, but a useful upper limit can be obtained, if the sensitivity of the system is known. The detector is a concave reflector of 10 m diameter and 7.3 m focal length, with 19

independent photomultipliers at the focus. It is located at the F. L. Whipple Observatory, Mount Hopkins, Arizona and technical details have been given elsewhere¹⁵. In this case it was lowered to point, almost horizontally, at a neighbouring mountain, Mt Wrightson, so that the background rate was small.

The interaction length for the catalysis process is $\lambda = 4,300 \beta / \sigma_0 \rho$ where σ_0 is a constant set at 1 for a strong interaction cross-section and ρ is the density of the medium in g cm^{-3} . Our system was sensitive to values of β in the region $\sim 2 \times 10^{-5}$, where the interaction length in air would be ~ 60 cm. A lower limit for β is set at about 10^{-5} by acceleration of monopoles in the Earth's gravitational field; an upper limit is set at about 5×10^{-5} by the increasing interaction length.

Our sensitive volume was approximately cylindrical in shape with a length of ~ 625 m, diameter 23 m, and surface area 45,000 m^2 . A slow monopole traversing this volume would cause typically 50 decays. The signal would come from the positron which is expected in the dominant decay mode, and to a lesser degree, from each of the γ rays produced in the neutral pion decay. The Cerenkov light would be emitted into three cones each of whose angular diameters would be determined both by Coulomb scattering and by the Cerenkov angle to about 3.4° half angle (0.06 radians). The probability of detecting one of the 50 decays would be 0.088 and hence of detecting two decays 7.7×10^{-3} . These would occur within 5 ms of the entry of the monopole into the observation volume.

The recording system had a dead time of 120 μs after a master pulse. To obtain some information on events within the dead time, a register was provided to detect pulses occurring between 1.5 and 85 μs after a master pulse. A master pulse was generated if three of the 19 photomultipliers were struck within 400 ns, and the incident light intensity required was 10 photons m^{-2} . The system was operated for 30 h, with a master pulse rate of $\sim 0.1 \text{ s}^{-1}$; 24 pairs of master pulses occurring within 5 ms were observed, a number consistent with random expectation. At the 90% confidence level this gave a limit of 28 per day. Four of these pairs also had events within the subsidiary register. They all occurred on the same night, and were probably caused by distant lightning, but must be included in the total since they satisfy the selection criteria.

The solid angle for incident monopoles on the selection volume is 4π so the limit F is given by

$$F = \frac{28}{4\pi \times 45,000 \times 7.7 \times 10^{-3}} \\ = 6.4 \times 10^{-3} \text{ m}^{-2} \text{sr}^{-1} \text{day}^{-1}$$

for β ranging from 10^{-5} to 5×10^{-5} and assuming $\sigma_0 = 1$. In c.g.s. units this corresponds to a value of $7.4 \times$

$10^{-12} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1}$. If σ_0 is <1 the probability of detection would be reduced, and the flux limit increased, by a factor approximately equal to $1/\sigma_0^2$.

The limit for $\sigma_0 = 1$ is about 15 times lower than that of Cabrera, but with a restricted range of β , and is comparable with experimental limits set at higher velocities. As it is considerably higher than values calculated from astrophysical arguments, its usefulness depends on the controversial effect of clustering around the Sun⁸. The mass of the monopole, which could be as high as 10^{19} GeV , is a significant factor in this⁷.

We thank Mr K. Harris and Dr P. K. MacKeown for help with observations. The work was supported in part by a grant from the National Board for Science and Technology of Ireland, by the United States Department of Energy, the Smithsonian Scholarly Studies Fund and the Kajima Foundation.

Received 3 May; accepted 28 June 1983.

1. Rubakov, V. A. *J.E.T.P. Lett.* **33**, 644–646 (1981).
2. Callan, C. G. *Nucl. Phys.* **B212**, 391–400 (1983).
3. Wilczek, F. A. *Phys. Rev. Lett.* **48**, 1146–1149 (1982).
4. Ellis, J., Nanopoulos, D. V. & Olive, K. A. Preprint TH3323, 1–9 (CERN 1982).
5. Kolb, E. W., Colgate, S. A. & Harvey, J. A. *Phys. Rev. Lett.* **49**, 1373–1375 (1982).
6. Zeldovich, Ya. B. & Khlopov, M. Yu. *Phys. Lett.* **79B**, 239–241 (1978).
7. Bais, F. A., Ellis, J., Nanopoulos, D. V. & Olive, K. A. Preprint TH3383, 1–37 (CERN 1982).
8. Dimopoulos, S., Glashow, S. L., Purcell, E. M. & Wilczek, F. A. *Nature* **298**, 824–825 (1982).
9. Cabrera, B. *Phys. Rev. Lett.* **48**, 1378–1381 (1982).
10. Cabrera, B. *Proc. Racine Conf. Magnetic Monopoles* (Plenum, New York, 1983).
11. Mashimo, T., Kawagoe, K. & Koshiba, M. *J. phys. Soc. Jap.* **51**, 3067–3068 (1982).
12. Bartelt, J. *et al.*, *Phys. Rev. Lett.* **50**, 655–658 (1983).
13. Groom, D., Loh, E., Nelson, H. N. & Ritson, D. M. *Phys. Rev. Lett.* **50**, 573–576 (1983).
14. Bosetti, P. C. *et al.* Preprint (1982).
15. Cawley, M. F. *et al.* *Proc. int. Workshop on High-Energy γ -ray Astronomy*, Ootacamund, 292–294 (1982).

Magnetic focusing in the Sco X-1 radio source

A. Achterberg, R. D. Blandford & P. Goldreich

Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125, USA

There has been much theoretical discussion of the confinement of the radio jets associated with extragalactic radio sources¹. There are now several examples of sources where the minimum pressure inferred in the jets appears to exceed the external gas pressure, which suggests that magnetic pinching may be playing an important part in the confinement^{2–4}. This issue has been highlighted by recent remarkable observations, using the Very Large Array, of the radio lobes associated with the galactic X-ray source Sco X-1 by Fomalont *et al.*⁵. We argue here that these results necessitate magnetic focusing and that this also strengthens the case for magnetic focusing in the extragalactic sources. We show that a sufficient overpressure with respect to the ambient interstellar medium can be achieved if the radius of the jet is reduced by a factor 10–100, and the converging flow becomes dissipative. The radio lobes which form at that point quickly die out again due to electron expansion losses in the now rapidly diverging jet. On the basis of this model it is predicted that the Faraday rotation changes sign across the radio lobes, and that the shape of the lobes should be conical with the apex pointing towards Sco X-1.

The observations, summarized in Table 1, show that the northern lobe is surprisingly compact and that the minimum internal pressure exceeds that of the ambient interstellar medium by a factor $\sim 10^6$. Furthermore, the lobe appears to be moving away from Sco X-1 slower than 30 km s^{-1} . Ram-pressure confinement of the lobe would require an external density in excess of $10^{-21} \text{ g cm}^{-3}$ which is unreasonably large. Inertial or self-gravitational confinement of the lobe would likewise require excessive mass $\geq 10^{-3} M_\odot$ which would have easily been observed. We therefore believe that the lobe is

Table 1 Observational parameters

Distance Sco X-1 = 500 pc
Radio luminosity lobes $L_r = 10^{29} \text{ erg s}^{-1}$
Diameter radio lobes $R = 4.8 \times 10^{-4} \text{ pc} = 1.5 \times 10^{15} \text{ cm}$
Separation of radio lobes Sco X-1 $D = 0.2 \text{ pc} = 7 \times 10^{17} \text{ cm}$
Equipartition field $B = 5 \times 10^{-3} \text{ G}$
Minimum energy density $U_{\min} = 2 \times 10^{-6} \text{ erg cm}^{-3}$

probably not confined along the source axis and instead must result from the focusing of a jet emerging from Sco X-1.

This view differs in two ways from that advanced by Fomalont *et al.* who suggested that the radio hotspots are formed in a high-density region. First, the radio components are located at a fixed distance from the origin of the jet, presumably the accretion disk powering the X-ray source Sco X-1. Second, the finite extent of the observable radio-lobes is now determined laterally by the radius of the jet at the constriction point, and along the jet axis by the rapid expansion of the jet behind the internal shocks, which causes the electrons to lose energy rapidly by expansion losses leading to a rapid decrease in surface-brightness.

Jets may be confined transversely either by gas pressure or by magnetic stresses. Gas-pressure confinement cannot cause significant focusing because the convergence speed of an unmagnetized jet towards its axis can never greatly exceed the internal sound speed. A more rapid convergence would give rise to internal (Mach) shocks whose dissipation would increase the internal jet entropy, inhibiting further compression (except in the artificial case in which the fluid is confined to the walls of the jet and the interior is evacuated)⁶. If the internal gas pressure follows an adiabat $P \propto R^{-2\gamma}$ with $\frac{4}{3} \leq \gamma \leq \frac{5}{3}$, the reduction in the jet radius is restricted to a factor of ~ 2 , as was found in a numerical calculation by Sanders⁷.

However, if the jet contains toroidal magnetic field (B_ϕ in cylindrical coordinates), the flow can converge at a speed comparable with the Alfvén speed by pinching forces. If the jet velocity V is super-Alfvénic, we expect that B_ϕ will scale inversely with the jet radius R , thus the Alfvén speed $a \equiv B_\phi / \sqrt{4\pi\rho}$ will remain roughly constant as the jet converges. The focal length from the point of maximum jet radius $R_{\max}$ is then $\frac{1}{2}D \sim R_{\max}(V/a)$.

The magnetic stress on the surface of the jet increases $\propto R^{-2}$. We assume that the return current needed to balance the current in the jet flows at a large distance $r = R_c$ from the jet. A toroidal magnetic field $\propto 1/r$ extends out to radius R_c where the magnetic stress is small enough to be balanced by the ambient interstellar pressure (see Fig. 1). This magnetic flux has presumably convected out of the jet into a cocoon of back-streaming material formed at its head.

We can make a simple model of a stationary converging flow under the influence of pinching forces by ignoring poloidal field, gas pressure and azimuthal motion, all of which inhibit focusing of the jet. The current then flows along streamlines and there is a simple self-similar solution of the equations of continuity, flux freezing and motion. This requires that the gas density be constant across the jet and that the field strength and the velocity towards the jet axis increase linearly with distance across the jet. The change in jet radius $\dot{R} \equiv V dR/dz$ can be found by balancing the centrifugal force $F_c \approx \rho V^2 d^2R/dz^2$ with the Lorentz force $F_l = -[B_\phi(R)^2]/2\pi R$ on the flow-lines of the boundary of the jet⁴. The resulting equation can be written as:

$$\frac{1}{2}\dot{R}^2 = 2a^2(R) \ln(R_{\max}/R) \equiv -\Phi(R) \quad (1)$$

where $a(R) \equiv [B_\phi(R)]/\sqrt{4\pi\rho}$, the Alfvén speed at the jet boundary, is constant by virtue of the conservation of mass-flux and electrical current in the jet. Here $\Phi(R)$ denotes an effective potential.

We do not expect conditions at the origin of the jet to correspond to this similarity solution. However, as the transverse speed is generally not much greater than the Alfvén speed

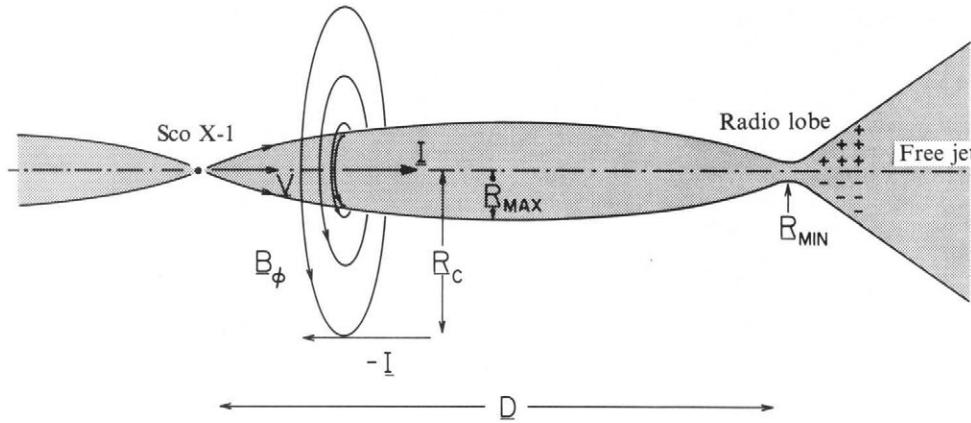


Fig. 1 Schematic representation of a magnetically focused jet in Sco X-1. The toroidal magnetic field B_ϕ reduces the jet width from its maximum value $R_{\max}$ to a value $R_{\min}$ at the hotspot where the jet is believed to become highly dissipative. The return current needed to balance the current carried by the jet flows at distance R_c of the jet axis. The + and - signs indicate the sign of the Faraday rotation in the lobes.

it should be possible for loops of magnetic flux to interchange in the effective radial gravitational field $\mathbf{g} = V^2(d^2R/dz^2)\hat{r}$ and thereby establish the magnetic and density profiles corresponding to the self-similar flow.

There remains the issue of stability. We assume that the velocity gradient across the jet is small except in a thin boundary layer, because any sizeable velocity gradient is expected to be unstable by the electromagnetic analogue of Richardson criterion⁸. We study the global stability of the jet by transforming into the co-moving frame and applying the energy principle⁹. Stability requires that the energy expended

$$\delta W = \frac{1}{2} \int [\xi \cdot \mathbf{g} \nabla \cdot (\rho_0 \xi) + \delta B^2/4\pi - \xi \cdot \mathbf{j}_0 \times \delta \mathbf{B}] dV \geq 0 \quad (2)$$

for all displacements ξ and associated magnetic perturbations $\delta \mathbf{B}$. We find that the self-similar flow is stable to arbitrary displacements that vanish at the jet boundary. The jet should therefore not be disrupted as long as the surrounding medium has sufficient inertia to resist transverse displacement. If the surface of the jet is free to move then the jet is probably unstable to Kelvin-Helmholtz modes. The overall persistence of the type of jet that we are proposing can only be determined by numerical calculation.

What then limits the focusing of the jet? There are two possibilities. We see from equation (1) that the Alfvén Mach number of the velocity of the jet boundary does increase, albeit very slowly. We assume that the flow can no longer stably adjust through the propagation of nonlinear Alfvén waves and oblique shocks will be formed which disrupt the flow when the jet radius has decreased by a factor $\sim 10^2$, which corresponds to an Alfvén Mach number $\approx 4-5$. The second possibility is that internal pressure, longitudinal field and toroidal motion inhibit collapse. We can incorporate all three effects as repulsive terms in the effective potential of equation (1) if we assume that they are self-similar too such that B_z , V_ϕ , $|\partial p/\partial r| \propto r$.

The effective potential $\Phi(R)$ then becomes:

$$\Phi(R) = 2a^2 \ln(R/R_{\max}) + \frac{1}{\gamma(\gamma-1)} s(R_{\max})^2 \left[\left(\frac{R}{R_{\max}} \right)^{2(1-\gamma)} - 1 \right] + \left(\frac{B_z^2}{8\pi\rho} + v_\phi^2(R_{\max}) \right) \left[\left(\frac{R}{R_{\max}} \right)^{-2} - 1 \right] \quad (3)$$

Here $s(R_{\max})$ is the internal sound speed of the jet at its maximum radius.

Motion in this potential yields periodic solutions in which the jet assumes a sausage-like shape along its axis. Such a behaviour was first seen by Chan and Henriksen⁴ in their numerical calculations of the structure of magnetically confined jets. We expect the oscillations to damp, or to disappear completely if near the constriction point some form of dissipation (like internal shocks) takes place, raising the internal entropy of the jet considerably. Note that it is impossible to confine the jet along its axis. So it continues either along its way forming secondary foci that are much less well confined and therefore invisible, or as a free jet which does not dissipate. The jet will

terminate when it has spread out sufficiently for its momentum flux to be balanced by the ram-pressure of the interstellar medium.

We have argued that the observed hotspots are located at the jet foci. If the transverse flow develops shocks when the jet has been focused to a radius $R_{\min}$ then the (average) relative kinetic energy associated with this transverse motion $\sim [2B^2(R_{\min})/3\pi] \ln(R_{\max}/R_{\min})$ can be made available for particle acceleration and field amplification. Note that this always gives an energy density close to the equipartition value. The requirement of a finite pressure P_e in the region external to the return current $r \geq R_c$ limits the field strength to $B_z^2/8\pi \leq P_e [R_c/R_{\min}]^2$. The internal pressure of the focus exceeds the pressure in the ambient interstellar medium by a factor $\approx 5.3 \times [R_c/R_{\min}]^2 \ln[R_{\max}/R_{\min}]$. Now $R_c \leq D$, so this factor is $\leq 5[D/R_{\min}]^2 \ln[R_{\max}/R_{\min}]$. Therefore it is possible to account for the large pressure $\approx 10^6 P_e$ inferred to be present in the focus, when the jet radius is reduced by a factor $R_{\max}/R_{\min} \approx 10-100$, depending on the value of R_c/D . In fact the pressure at the foci could be still larger if sufficient gas were entrained from the sheath at the focus to decelerate the jet and thereby dissipate some of the longitudinal kinetic energy which exceeds the energy associated with the convergence by an additional factor

$$\approx (V/a)^2 \ln(R_{\max}/R_{\min}) \sim \frac{D^2}{\pi R_{\max}^2} \ln(R_{\max}/R_{\min})$$

As there is substantial dissipation in the focus, the jet fluid will expand at the post-shock sound speed $\sim a(R_{\min})$. We assume that the electrons have a power-law distribution and expand adiabatically with the jet fluid. Since the magnetic field scales as R^{-1} , the surface brightness I_ν at a given frequency will scale as $^{10} I_\nu \propto R^{-(7\alpha+6)/3}$, where α is the spectral index. For a free jet one has $R \propto x$, where x is the distance from the focal point. The shape of the radio lobes when observed with high enough resolution should be roughly conical and extending away from Sco X-1 with highest intensity of emission concentrated near the peak of the cone.

There is a second prediction that can be made. If the foci are produced by the pinching action of toroidal magnetic field, and there is a sufficient density of plasma, then we would expect a gradient of Faraday-rotation depth across each radio component. The gradient should change sign on opposite sides of Sco X-1. In particular, the position of the foci may shift with the direction of linear polarization. If we use the self-similar solution, then the maximum rotation measure (RM) is

$$RM \approx \left(\frac{n_e}{1 \text{ cm}^{-3}} \right) \left(\frac{B(R_{\min})}{10^{-2} \text{ G}} \right) \left(\frac{R_{\min}}{10^{15} \text{ cm}} \right) \text{ rad m}^{-2}$$

If we substitute the field and size from Table 1, we find that the rotation measure satisfies $RM \approx 400(a/1,000 \text{ km s}^{-1})^{-2} \text{ rad m}^{-2}$ and should be detectable.

We have described a simple model for the magnetic focusing of the radio source in Sco X-1. The total power carried by the jet is estimated to be $\sim 10^{32} (V/10,000 \text{ km s}^{-1})(D/R_{\max})^2 \text{ erg}$

s^{-1} . This is small compared with the total X-ray luminosity of Sco X-1 ($L_X \approx 10^{37}$ erg s^{-1} for the assumed distance) but large compared with the radio luminosity. The current flowing along the jet is conserved and is $I \sim 3 \times 10^{13}$ A. This may originate from a centrifugally driven magnetohydrodynamic wind leaving a corona above an accretion disk. The associated magnetic field strength, if the Alfvén point is at radius R_A , is $\sim 10^4 (R_A/10^9 \text{ cm})G$ which is quite compatible with existing models of accretion disks in low-mass binaries.

Another galactic source in which some form of focusing appears to be occurring is SS433. Note that the angle subtended by the 'ears' of the associated radio source W50 at the central object is smaller than the angular width of the precession cone inferred from the kinematics of the moving emission lines¹¹. If magnetic focusing is substantiated in these galactic sources, then this strengthens the case for magnetic confinement in the extragalactic jets. Furthermore, it suggests the possibility of applying scaling arguments to infer the physical conditions in the accretion disks around massive black holes in active galactic nuclei.

We thank Ed Fomalont for bringing these observations to our attention.

Received 25 April; accepted 9 June 1983.

1. Miley, G. M. *A. Rev. Astr. Astrophys.* **18**, 165 (1980).
2. Benford, G. *Astrophys. J.* **247**, 792 (1981).
3. Pottasch, R. & Wardle, J. F. C. *Astrophys. J.* **84**, 707 (1979).
4. Chan, K. I. & Henriksen, R. N. *Astrophys. J.* **241**, 534 (1980).
5. Fomalont, E. B., Geldzahler, B. J., Hjellming, R. M. & Wade, G. M. *Astrophys. J. Lett.* (submitted).
6. Eichler, D. Preprint, University of Maryland, 1982).
7. Sanders, R. H. *Astrophys. J.* **266**, 73 (1983).
8. Chandrasekhar, S. *Hydrodynamic and Hydromagnetic Stability* (Cambridge University Press, 1961).
9. Bateman, G. *MHD Instabilities* (MIT Press, Cambridge, 1978).
10. Kellermann, K. I. & Pauliny-Toth, H. K. A. *Rev. Astr. Astrophys.* **19**, 373 (1981).
11. Geldzahler, B. J., Pauls, T. & Salter, C. J. *Astr. Astrophys.* **84**, 237 (1980).

Is the polarization of NGC1068 evidence for a non-thermal source?

I. S. McLean*, C. Aspin†, S. R. Heathcote†
& M. J. McCaughrean†

* Royal Observatory, and † Department of Astronomy,
University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK

NGC1068 is one of the brightest galaxies included by Seyfert¹ in his list of extragalactic objects having compact, luminous nuclei within which broad, high-excitation emission lines occur. It has been the subject of intensive studies at UV^{2,3}, optical³⁻⁶, IR⁷⁻¹² and radio wavelengths¹³⁻¹⁵. Unresolved questions concern the nature and relationship of the sources of the excess flux seen in the UV and IR, their connection with the collimated jets apparent in high-resolution radio maps and their association with the extended region responsible for the broad emission lines. A further question is the location of any dust and its role in modifying the optical and UV spectrum. We report here observations with two high-resolution optical spectropolarimeters which throw new light on these questions. From detailed structure found in the linear polarization spectrum of the nucleus we conclude that dilution by starlight modifies the polarization to an extent not previously appreciated. In fact, the polarization of the non-stellar flux in the optical and near IR is approximately independent of wavelength (as expected for synchrotron emission or electron scattering) with a direction orthogonal to that of the radio jets; such an arrangement is reminiscent of certain quasars and radio galaxies^{16,17}.

Optical polarization in NGC1068 was first reported by Walker¹⁸ who found that the observed polarization rose strongly into the blue. Subsequently, Visvanathan and Oke¹⁹ attempted to account for the observed flux and polarization as a superposi-

tion of an unpolarized stellar component and a non-stellar component with polarization independent of wavelength. However, it was not until the work of Angel *et al.*²⁰ that spectropolarimetry with enough resolution to separate the effects of lines and continuum was obtained. Those measurements had a variable resolution (20 Å at 3,300 Å; 160 Å at 6,500 Å) and were sufficient only to demonstrate general trends. Accurate broad-band polarimetry²⁰ revealed a weak (5°) rotation of position angle (θ) with wavelength (λ) across the optical region. In addition, a small component of circular polarization was detected ($\sim 0.2\%$). It was concluded^{20,21} that these results favoured a dust-scattering polarization mechanism.

Recently, we have observed NGC1068 using two new multi-channel spectropolarimeters. One of these is an imaging spectropolarimeter (ISP) with a charge coupled device (CCD), and the other is a Pockels cell polarimeter used in conjunction with the spectrograph and Image Photon Counting System (IPCS) at the Anglo-Australian Telescope. Both of these polarimeters (designed and built at the Royal Observatory Edinburgh) are described elsewhere^{22,23}.

The observations of NGC1068 reported here were made on the 3.9-m Anglo-Australian Telescope. Those with the CCD system were obtained on 29 August and measurements with the Pockels cell/IPCS were taken on 2 September 1982. The entrance apertures used were 1.7×2.2 arc s for the CCD and 2.0×2.5 arc s for the IPCS. Instrumental calibrations have been applied using standard techniques^{22,23}; the IPCS data are photon noise-limited whereas the CCD data contain some extra noise due to CCD defects and charge shifting losses. Figures 1 and 2 display the ISP and Pockels cell data, respectively.

One feature of the new data is that the $p(\lambda)$ curve is not at all smooth but exhibits a complex structure of bumps and dips which are not related to the emission lines which dominate the spectrum. Of particular interest is the region from $\lambda 4,500$ to $\lambda 5,300$ Å because this interval is covered by both instruments. By comparing Figs 1 and 2 it is quite evident that the broad area (or notch) of low polarization around 4,750 Å has been detected by both polarimeters. On re-examining the earlier low-resolution observations of Angel *et al.*²⁰ we noticed that the '4,750 Å notch' and several other ripples were indeed present but were apparently mistaken for noise.

There is also structure associated with the emission lines themselves. In our data, the [O III] lines at 4,959 and 5,007 Å are resolved. Across both lines there is a strong depolarization and a very large ($\sim 45^\circ$) rotation of θ . In the Pockels cell data (see Fig. 2) there is perhaps evidence for substructure in p and θ within the 5,007-Å line. There may also be structure in p associated with the [S II] lines and the [N II]/H α blend (Fig. 1). Because we observe the flux and polarization we can derive the polarization of the emission lines by themselves by interpolating for the continuum polarization in the line. Using the Pockels cell data for the strong 5,007 line the average intrinsic polarization of the emission is $p = (1.3 \pm 0.2)\%$ at $\theta = 151 \pm 4^\circ$. This result agrees with the values obtained by Angel *et al.* for aperture sizes of 2 and 3 arc s which are $(0.6 \pm 0.2)\%$ at $143 \pm 8^\circ$ and $(1.0 \pm 0.1)\%$ at $149 \pm 4^\circ$.

For the permitted lines the situation is ambiguous because of the complexity of the continuum $p(\lambda)$ curve. The H α line and [N II] lines are blended, H β and He II (4,686) lie in the 4,750 Å notch and H γ is weak. Nevertheless, it seems likely that part of the decrease in p within the 4,750-Å notch may be due to unpolarized He II and H β emission.

In addition to the continuum and line structure already mentioned, our data reveal another new feature, namely relatively narrow enhancements in p just redward of the Balmer emission lines. As the H β feature is recorded by both polarimeters we are confident that it is not an instrumental effect. The feature is broadened by binning in the Pockels cell data and by the instrumental profile in the CCD data. These enhancements in p do not correlate with any obvious intensity structure in the red wings of the Balmer lines but resemble displaced 'mirror images' of the emission lines themselves. The redshift of these

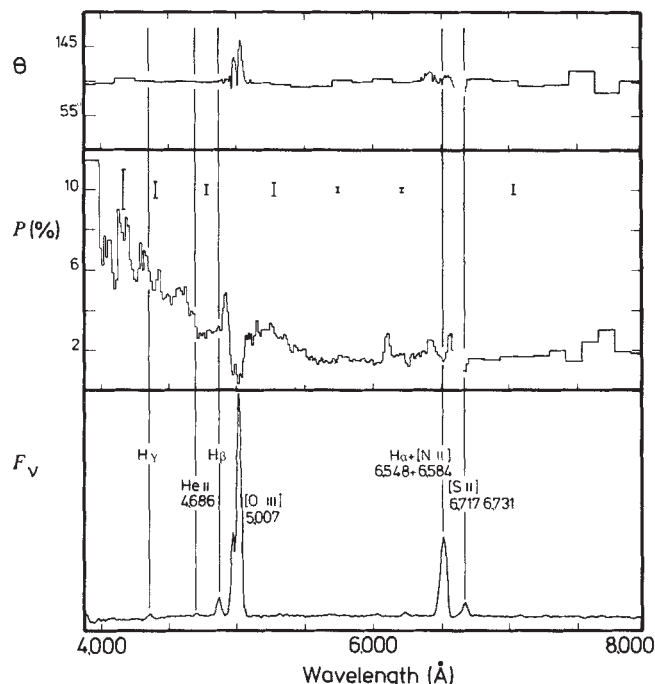


Fig. 1 CCD imaging spectropolarimeter data showing the position angle (θ), the percentage linear polarization (p) and flux (F_v) as a function of wavelength. Typical error bars per bin in p are shown; the error in θ is $28.65^\circ \sigma_p/p$. The bin width is $15 \text{ \AA}$ over most of the p data; the resolution is $\sim 30 \text{ \AA}$. A small segment around $\lambda 6,700 \text{ \AA}$ has been deleted due to a group of bad pixels.

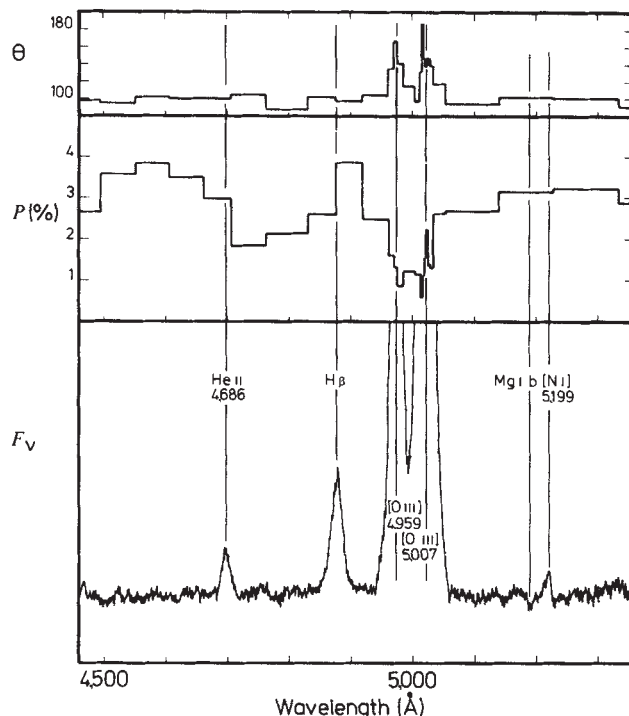


Fig. 2 Pockels cell/IPCS data. Position angle (θ), percentage linear polarization (p) and flux (F_v) are shown as a function of wavelength. The data have been binned such that the expected photon noise error in p is 0.5% outside the $[O III]$ lines and 0.4% within the lines. At the full resolution a single bin corresponds to $0.5 \text{ \AA}$.

features relative to the Balmer line cores is $\sim 50 \text{ \AA}$ (that is $\sim 3,000 \text{ km s}^{-1}$ at $H\beta$).

To interpret these observations we first need to consider qualitatively the continuum polarization. Any underlying polarized continuum arising in the nucleus of NGC1068 must inevitably be diluted by the integrated light of stars in that part of the galaxy in the line of sight. Therefore, the observed polarization should show largest values (least dilution) at wavelengths corresponding to stellar absorption lines. For a normal spiral galaxy the spectrum is dominated by solar and late-type stars, and there are absorptions due to H and K bands of Ca (the $4,000 \text{ \AA}$ break), the G band of CH ($4,300 \text{ \AA}$), the $Mg I b$ ($5,174 \text{ \AA}$) feature, the Balmer lines of hydrogen and molecular bands of several species of molecules. In addition, the decline of the stellar spectra towards short wavelengths will result in less dilution and therefore more polarization in the blue.

From observations of the galaxy itself¹⁸ the position angle of the rotation axis is 145° . This is close to the direction of the polarization of the $[O III]$ lines. Polarization due to transmission through grains aligned by a galactic magnetic field is expected to produce polarization vectors parallel to the galactic plane (that is, $\theta \sim 55^\circ$ not 145°). On the other hand, polarization of the $[O III]$ lines could be due to scattering into our line-of-sight by grains on the outskirts (farthest from the nucleus) of the emitting regions. On average, these emitting regions seem to be aligned with the projection of the major axis of the galaxy. Because dilution by galactic light will affect the emission lines less than the neighbouring continuum the permitted lines must in fact be substantially less polarized than the continuum, contrary to the findings of Angel *et al.* The absence of any obvious change in θ does, however, indicate that they do not share the polarization of the forbidden lines.

Since dilution leaves the plane of polarization invariant it cannot account for the weak wavelength dependence of θ revealed by broadband observations unless the galaxy light is itself slightly polarized at a different angle from the underlying source. Such polarization can readily occur by passage through the interstellar medium of NGC1068. Because the galactic flux

is weakest in the far blue, the observed position angle of 102° should correspond most nearly to that of the non-stellar component. A direction of 102° is almost orthogonal to the projected direction of the radio jet, once allowance is made for its curvature^{14,15}. We estimate the jet to lie in position angle $\sim 11^\circ$ in the optical nucleus whereas the elongated central radio component itself would give $\sim 20^\circ$; this orthogonal geometry also pertains further out in the jet as demonstrated by 15-GHz polarization data¹⁴. If the polarized emission is interpreted as being due to the synchrotron process then the magnetic field is aligned along the jet.

The enhancements in p associated with the Balmer lines are puzzling. One explanation may be that they correspond to strong redshifted absorption lines. At the wavelengths of these absorption lines dilution would be considerably reduced and so the polarization would tend to rise. As far as we know, no set of redshifted absorption lines has been detected in the spectrum. Alternatively, they may be due to cloud-cloud scattering, perhaps in high-velocity jet material. For example, a photon emitted from a cloud moving in one direction could be scattered by a cloud moving in the opposite direction and thus suffer a shift in wavelength. The dispersion of velocities observed in NGC1068 is almost high enough to be consistent with such a model ($\sim 2,000 \text{ km s}^{-1}$).

To quantify the idea that dilution has a crucial role we have constructed a simple, semi-empirical model for the continuum. The underlying non-stellar (nuclear) component was taken to be a power law $F_v \propto \nu^{-\alpha}$ with an intrinsic polarization p_{ns} independent of λ (as is appropriate for optically-thin synchrotron emission from a homogeneous source or for electron scattering) at a position angle perpendicular to the radio jet so that $100^\circ < \theta_{ns} < 120^\circ$. The spectrum of the galactic component was taken to be the same as that of the nucleus of M31²⁴ (which, like NGC1068, is of type Sb). The interstellar polarization of the galactic component was assumed to follow Serkowski's empirical relationship²⁵, which is characterized by two parameters λ_{max} and p_{max} (the wavelength of maximum polarization and its value). We adopt a position angle for the polarization

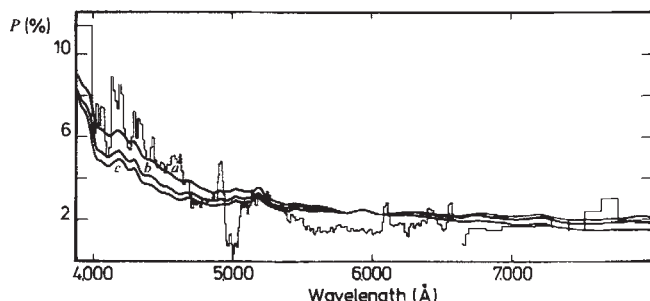


Fig. 3 Representative results from the continuum polarization model described in the text. For all three $p(\lambda)$ curves the interstellar contribution in NGC1068 is given by $p_{\max} = 0.6\%$, $\lambda_{\max} = 5,000 \text{ Å}$, $\theta_{\text{gal}} = 55^\circ$; the position angle of the non-stellar component is $\theta_{\text{ns}} = 102^\circ$; the template galaxy is M31 and both the non-stellar and the galactic components suffer an extinction $A_v = 0.4 \text{ mag}$. The remaining parameters for each curve are as follows. Model a: $\alpha = 0.0$, $F_{\text{ns}} = 0.23$, $P_{\text{ns}} = 12\%$; Model b: $\alpha = 1.0$, $F_{\text{ns}} = 0.20$, $P_{\text{ns}} = 13\%$; Model c: $\alpha = 1.5$, $F_{\text{ns}} = 0.166$, $P_{\text{ns}} = 15\%$. Model a gives the best fit to our optical data while model c reproduces the IR results.

of the galactic component $\theta_{\text{gal}} \sim 55^\circ$ as expected for extinction by grains aligned by a magnetic field lying parallel to the galactic plane of NGC1068. To specify the relative contributions of the two components we define F_{ns} as the fraction of total light contributed by the non-stellar component at $5,500 \text{ Å}$ within a circular aperture 2 arc s in diameter. Either or both components may be reddened; the wavelength dependence of the extinction (A_v) is assumed to follow the average relationship for our galaxy²⁶. Figure 3 presents the results of some representative models superimposed on the ISP data.

Values of $\lambda_{\max}$ and $p_{\max}$ are constrained by the weak $\theta(\lambda)$ dependence in the continuum and F_{ns} is (for a given α) constrained by the need to match the observed spectrum. The principal free parameters are therefore α and p_{ns} . Acceptable fits to our data have $0 < \alpha < 1.5$; larger values produce too small a variation in p between 4,000 and 5,000 Å while smaller values cause too rapid a decline into the red. To constrain the model parameters further it is desirable to use IR observations. It seems^{12,27} that the trends seen in the optical polarization reverse beyond $1 \mu\text{m}$ so that p increases to a broad maximum ($\sim 4.2\%$ in a 4 arc s aperture) near $2.2 \mu\text{m}$, while θ rotates back to larger angles. This is also a dilution effect as normal galaxies peak around $1.6 \mu\text{m}$. To produce such an upturn at all, however, requires values of α in the upper part of the range given above. In fact, the model having $\alpha = 1.5$ yields $p = 1.6\%$, 2.4% and 4.4% at 1.25 , 1.65 and $2.2 \mu\text{m}$ respectively, in reasonable accord with the results of Lebofsky *et al.*¹² but is not so satisfactory in the optical. A value of $\alpha \sim 1.0$ is a plausible compromise. The departure from the model in the $5,200$ – $6,500 \text{ Å}$ range may indicate that the synchrotron source is actually inhomogeneous. In common with other studies^{2–4}, the values of A_v in our model are very much smaller than those found for the emission line regions themselves^{3,12,28,29}. This would seem to imply that much of the dust lies within the emission-line clouds.

The values of α and F_{ns} found here are in good agreement with those obtained by Koski⁴ from modelling of the intensity spectrum alone. M. Ward (personal communication) has estimated $F_{\text{ns}} \sim 30\%$ from published aperture photometry. Recent estimates based on a comparison of the strengths of stellar absorption lines in the blue with those in normal galaxies^{30,31} imply $F_{\text{ns}} \sim 70\%$. With such a large value of F_{ns} neither a simple synchrotron model nor a pure scattering model is tenable but an inhomogeneous synchrotron source cannot be ruled out.

A model invoking pure dilution of synchrotron emission must account for the observation of a small but significant component of circular polarization. Normal synchrotron emission could only account for this circular polarization if the magnetic field were very strong, $\sim 10^3 \text{ G}$. Two alternative explanations can be advanced, namely, multiple scattering in dust clouds (which

would also give a large linear polarization due to scattering) or single scattering of a highly linearly polarized beam through a medium of aligned grains^{32,33}. The latter mechanism may be efficient in NGC1068 because the difference of position angles between synchrotron polarization and grain alignment is near to the optimum (that is, 45°).

We thank the staff of the AAO for support, and our colleagues Malcolm Smith, Adrian Webster and Ray Wolstencroft, Martin Ward and David Axon for their comments. This work was supported by the SERC.

Received 26 April; accepted 16 June 1983.

- Seyfert, C. K. *Astrophys. J.* **97**, 28–40 (1943).
- Boksenberg, A. *et al. Nature* **275**, 404–414 (1978).
- Neugebauer, G. *et al. Astrophys. J.* **238**, 502–509 (1980).
- Koski, A. T. *Astrophys. J.* **223**, 56–73 (1978).
- Shields, G. A. & Oke, J. B. *Astrophys. J.* **197**, 5–16 (1975).
- Penston, M. V., Penston, M. J., Selmes, R. A., Becklin, E. E. & Neugebauer, G. *Mon. Not. R. astr. Soc.* **169**, 357–393 (1974).
- Becklin, E. E., Mathews, K., Neugebauer, G. & Wynn-Williams, C. G. *Astrophys. J. Lett.* **186**, L69–L72 (1973).
- Telesco, C. M., Becklin, E. E. & Wynn-Williams, C. G. *Astrophys. J. Lett.* **241**, L69–L72 (1980).
- Hildebrand, R. H. *et al. Astrophys. J.* **216**, 698–705 (1977).
- Cutri, R. M. *et al. Astrophys. J.* **245**, 818–828 (1981).
- McCarthy, D. W., Low, F. J., Kleinmann, S. G. & Gillett, F. C. *Astrophys. J. Lett.* **257**, L7–L11 (1982).
- Lebofsky, M. J., Rieke, G. H. & Kemp, J. C. *Astrophys. J.* **222**, 95–99 (1978).
- de Bruyn, A. G. & Willis, A. G. *Astr. Astrophys.* **33**, 351–356 (1974).
- Wilson, A. S. & Ulvestad, J. S. *Astrophys. J.* **263**, 576–594 (1982).
- Pedlar, A., Boller, R. V., Spencer, R. E. & Stewart, O. J. *Mon. Not. R. astr. Soc.* **202**, 647–656 (1983).
- Stockman, H. S. & Angel, J. R. P. *Astrophys. J. Lett.* **220**, L67–L71 (1978).
- Antonucci, R. R. J. *Nature* **299**, 605–606 (1983).
- Walker, M. F. *Astrophys. J.* **151**, 71–97 (1968).
- Visvanathan, N. & Oke, J. B. *Astrophys. J. Lett.* **152**, L165–L168 (1978).
- Angel, J. R. P., Stockman, H. S., Woolf, N. J., Beaver, E. A. & Martin, P. G. *Astrophys. J. Lett.* **204**, L5–L9 (1976).
- Nandy, K. & Wolstencroft, R. D. *Nature* **225**, 621 (1970).
- McLean, I. S., Cormack, W. A., Herd, J. T. & Aspin, C. *Proc. Soc. Photo-opt. Instrum. Engng* **290**, 155–200 (1981).
- McLean, I. S., Heathcote, S. R., Paterson, M. J., Fordham, J. & Shortridge, K. *Mon. Not. R. astr. Soc.* (submitted).
- Oke, J. B. & Schwarzschild, M. *Astrophys. J.* **198**, 63–70 (1975).
- Serkowski, K. in *Interstellar Dust and Related Topics* (eds van de Hulst, H. C. & Greenberg, J. M.) 145–159 (Reidel, Dordrecht, 1973).
- Savage, B. D. & Mathis, J. S. A. *Rev. Astr. Astrophys.* **17**, 73–111 (1979).
- Knacke, R. F. & Capps, R. W. *Astrophys. J. Lett.* **192**, L19–L22 (1974).
- Malkan, M. A. *Astrophys. J. Lett.* **264**, L1–L6 (1983).
- Gaskell, C. M. *Publ. astr. Soc. Pacif.* **94**, 891–893 (1982).
- Malkan, M. A. & Oke, J. B. *Astrophys. J.* **265**, 92–106 (1983).
- Malkan, M. A. & Filippenko, A. V. *Astrophys. J.* (submitted).
- Martin, P. G., Illing, R. & Angel, J. R. P. *Mon. Not. R. astr. Soc.* **159**, 191–201 (1972).
- Lonsdale, C. J., Dyck, H. M., Capps, R. W. & Wolstencroft, R. D. *Astrophys. J. Lett.* **238**, L31–L34 (1980).

Three-dimensional chemical waves in the Belousov–Zhabotinskii reaction

Brian J. Welsh, Jagannathan Gomatam & Arthur E. Burgess

Departments of Mathematics and Chemistry, Glasgow College of Technology, Cowcaddens Road, Glasgow G4 0BA, UK

A simply-connected region of the Belousov–Zhabotinskii (B–Z) reaction medium may produce a variety of three-dimensional wave forms of unusual geometry. The reaction–diffusion hypothesis even in its linear formulation supports iso-concentration surfaces which include helical surfaces and toroidal scroll waves with twists¹. A variety of other wave forms has been proposed by Winfree and Strogatz². Experimental difficulties associated with observations in three dimensions have precluded confirmation of these waves *in situ*. We now present direct experimental evidence of evolving three-dimensional chemical waves. Our design is easy to implement, lends itself to time-delay photography and is essentially free of difficulties pointed out earlier³. The dominant wave forms are toroidal scrolls^{1,3–5}; other wave forms observed include the sphere-like structure predicted recently².

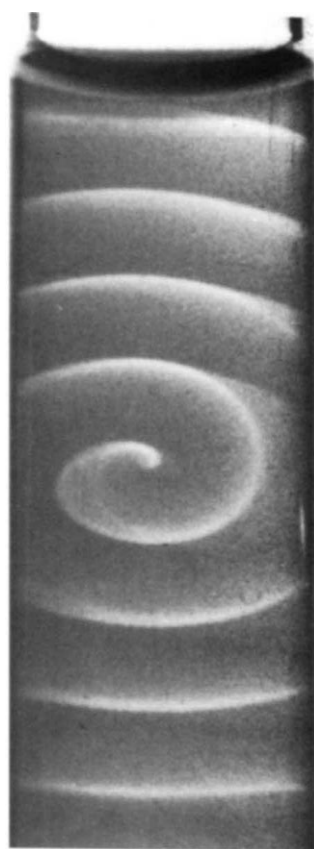


Fig. 1 A simple scroll wave⁵ seen in cross-section in a test tube of diameter 10 mm. Close examination of this and the other figures shows no distortion from convection currents nor is there any evidence of gas bubbles.

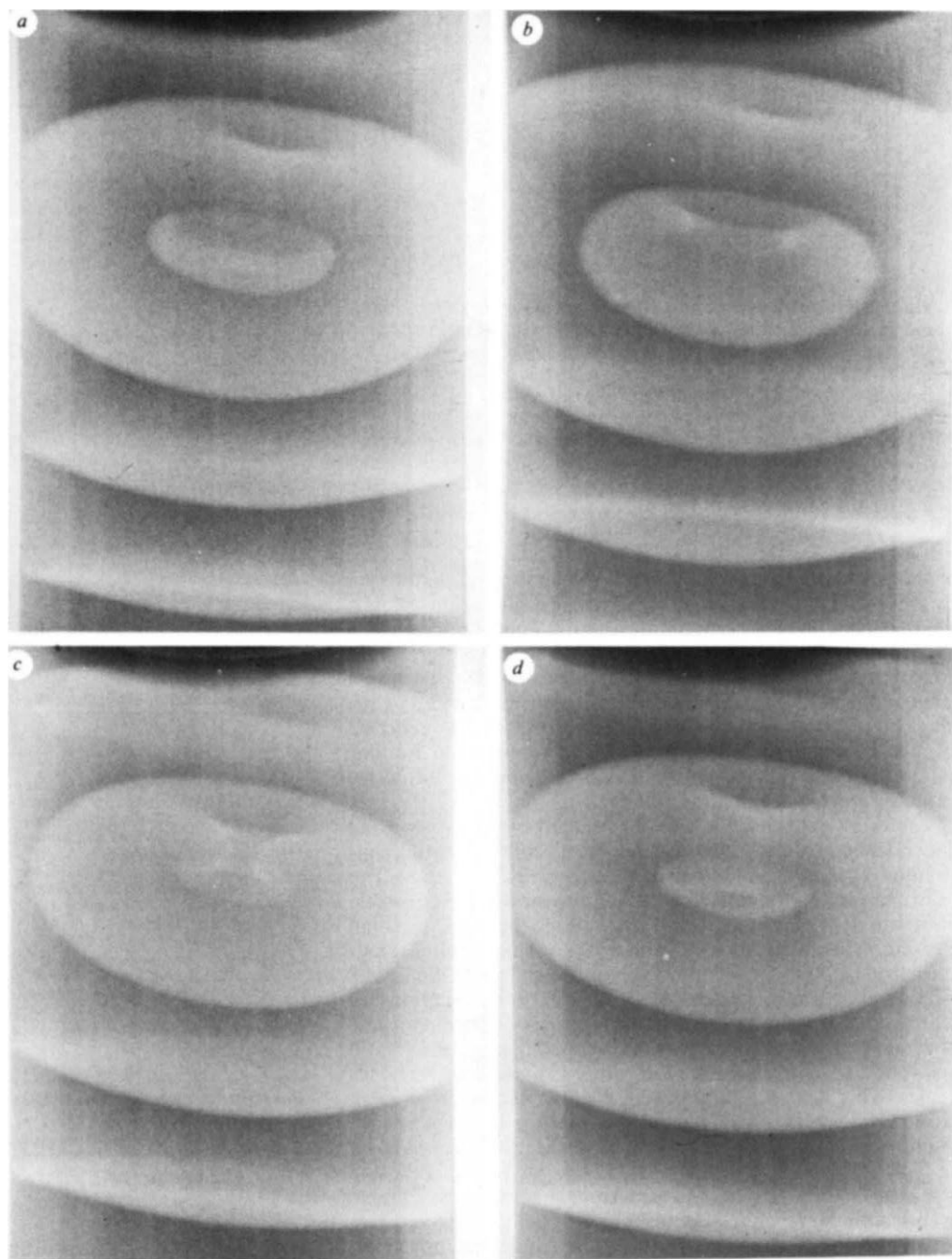


Fig. 2 *a-d*, A subsequent stage in the evolution of a simple toroidal scroll into a sphere-like structure. This appears to evolve from and rotate around a central core filament. Photographs were taken at ~25-s intervals and covered one period of 115 s throughout the evolution of the structure. The reaction occurs in a glass test tube (diameter 10 mm), the developing pattern being annihilated when it came into contact with the glass wall. Eventually iso-concentration surfaces form a series of parallel shells on either side of the central torus.

When attempts are made to observe the evolution of three-dimensional waves in the B-Z reaction the slightest perturbation, thermal or mechanical, disrupts the developing patterns. For example, the reaction generates bubbles of carbon dioxide gas and as these bubbles rise to the surface they destroy any patterns through which they pass. Fluid motion has a similar effect and since convection currents in deep solutions are almost inevitable, wave surfaces in media having a depth of even a few millimetres are often grossly distorted.

Winfree⁴ developed an ingenious technique to overcome some of these problems. Three-dimensional patterns were obtained by initiating a wave in a stack of Millipore filters soaked in the reaction solution. When the wave had evolved, the filters were separated and a record of the wave structure

obtained. However, with this technique it is not possible to observe three-dimensional waves as they evolve and the structure still needs to be deduced from two-dimensional sections of the Millipore stack. Furthermore useful patterns are obtained only in shallow layers, the overall depth being of the order of a millimetre or so. Indeed to our knowledge experimental conditions for watching three-dimensional chemical waves evolve undistorted and to a level of definition suitable to photography has not yet been reported. The present procedure allows direct observation and uses the B-Z reagents in a straightforward manner. Typically the reaction mixture is prepared by adding 1 ml of 1.5 M sulphuric acid to 750 ml of water and dissolving in it 1.74 g of cerous nitrate. To 10 ml of this solution is added 10 ml of each of the following solutions:

0.35 M potassium bromate, 1.5 M sulphuric acid and 1.2 M malonic acid. The solution in a cylinder of diameter 27 mm is stirred smoothly and steadily with a magnetic follower for 5 min. Then 0.6 ml of the indicator dye ferroin (phenanthroline 25 mM) is added and stirring is continued for a further 35 min at a room temperature of $\sim 22^\circ\text{C}$.

Stirring eliminates phase gradients, although it is known that the reaction is sensitive to oxygen transferred from the air⁶. Consequently a cylinder of small diameter is used and formation of a deep vortex is avoided so that the ingress of atmospheric oxygen is restricted. Next, 3-ml portions of the prepared reagent are poured slowly into several test tubes of internal diameter 8 mm and length 75 mm. The length of test tube is arbitrary but its diameter is crucial. It has been found that in a test tube of diameter < 15 mm distortion of patterns by convection currents is negligible. Rubber stoppers are inserted into the test tubes which are then set against a background of diffuse light. Since it is essential that a constant temperature be maintained in the vicinity of each tube any thermal current from draughts or lightbox must be deflected.

After a short period, while the system remains quiescent, patterns will start to evolve in the tubes. Most structures formed are complicated and difficult to identify, but usually after an interval a dominant wave form emerges. Since not every container will support waves or easily recognizable structures an array of at least 10 test tubes is recommended.

The system has the advantage of being insensitive to small mechanical perturbations. In particular, test tubes can be slowly rotated and a pattern which appears disguised in one perspective may be more easily recognizable in another. Moreover, an overall perspective of the evolving structure can be gained. Patterns can be observed continually for up to 5 h. This depends on the initial concentrations chosen and especially on that of the ferroin. Some test tubes contain a series of parallel bands or simple scroll waves (see Fig. 1). By lightly tapping these tubes a variety of other structures can be created. From the initial disorder, wave forms of unusual geometry sometimes emerge.

Sealing the test tubes is crucial to this design for two reasons. First, the pressure created by inserting rubber bungs greatly reduces, indeed almost eliminates, the formation of gas bubbles.

If test tubes are not initially sealed under pressure, gas bubbles will form in abundance on the sides of the test tubes. Second, rubber bungs exclude the ingress of atmospheric oxygen which has a disruptive effect on patterns formed.

Gas bubbles still occur in some of the test tubes ruining any incipient patterns. This problem can be greatly reduced by pouring the reagent slowly into the test tubes. It has been found that steeping the test tubes in deionized water for an hour before use reduces the number of nucleation sites on which bubbles form.

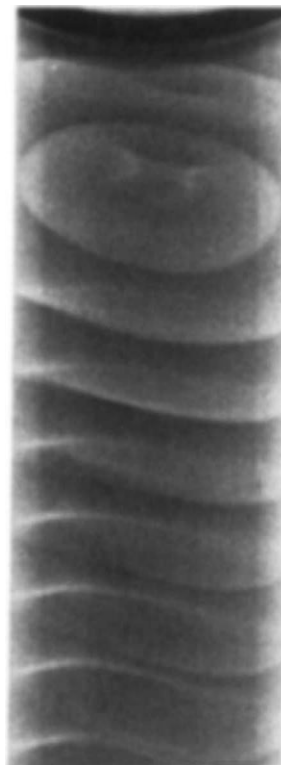


Fig. 3 An overall view of the structure in Fig. 2.

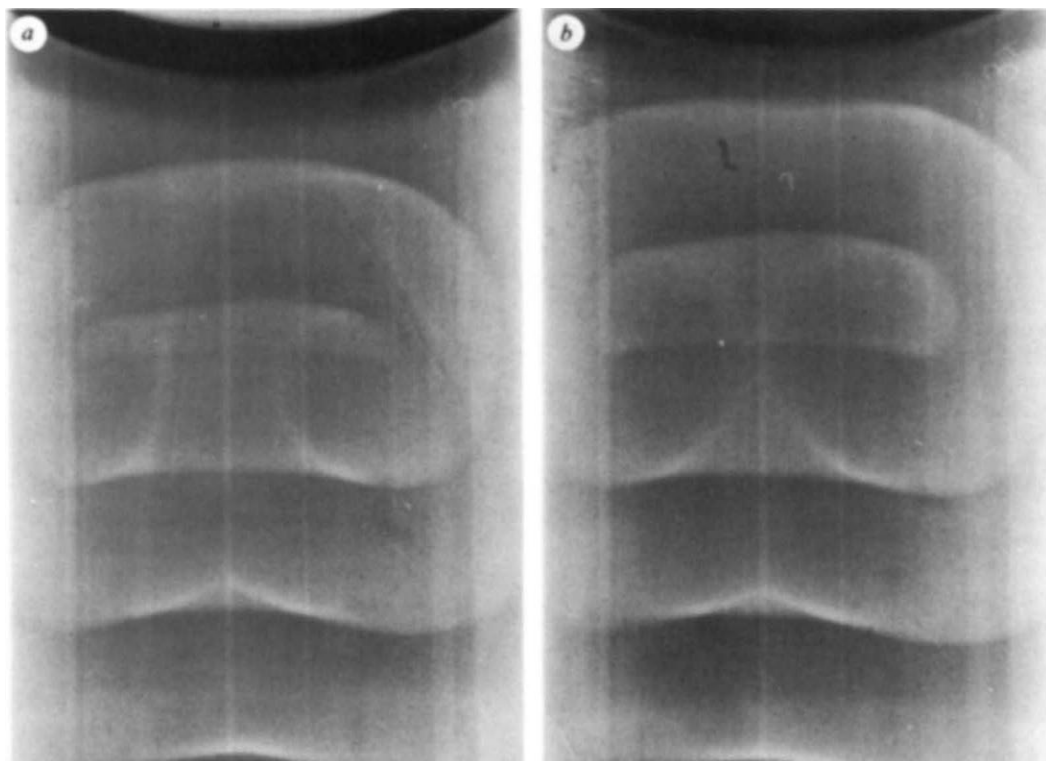


Fig. 4 A developing scroll wave⁵ sheared by the sides of the test tube. Two oppositely wound spiral waves are seen in cross-section to emerge from the core filament.

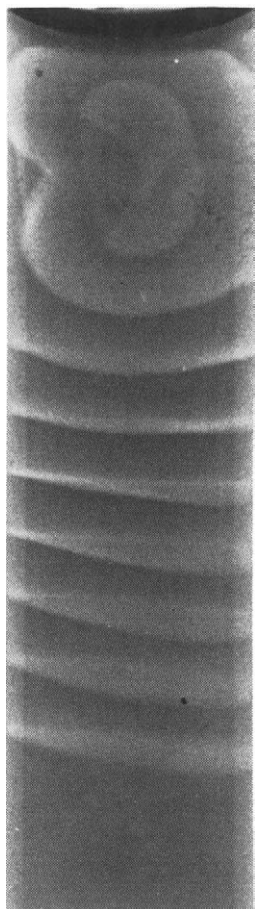


Fig. 5 An unusual type of spheroidal wave.

Figure 2 shows a subsequent stage in the evolution of a simple toroidal scroll into a sphere-like structure. This appears to evolve from and rotate around a central core filament. Photographs were taken at ~25 s intervals and covered one period of 115 s throughout the evolution of the structure. The reaction occurred in a glass test tube of internal diameter 8 mm, with the developing pattern being annihilated when it came into contact with the glass wall. Eventually iso-concentration surfaces formed a series of parallel shells on either side of the central torus (see Fig. 3). Over a time interval of ~60 min the major radius of the central torus shown in Fig. 2 decreases and eventually the torus itself disappears. A scroll wave sheared by the sides of the test tube is shown in Fig. 4. Two oppositely-wound spiral waves are seen in cross-section to emerge from the core filament. Two kinds of patterns were most frequently observed: simple scrolls and sphere-like structures. Some of the wave surfaces as they evolve assume a variety of intermediate forms of unusual geometry that are not easy to describe. An example of this is given in Fig. 5.

This experimental design has substantial advantages. The volatile, toxic by-products produced by the reaction are contained and as discussed earlier, freedom from disruptive effects enables one to observe and record the long-term history of the evolving patterns.

These features recommend the design for further experiments involving, for example, multiply-connected domains. Our observations raise the important question of whether the dimension of the vessel imposes a limitation on the geometric diversity of potential wave forms.

We thank Dr N. G. Meadows for his interest in the project and Gordon Ridley and Douglas Neil for assistance with the photography.

Received 22 March; accepted 3 June 1983.

1. Gomati, J. *J. Phys. A* **15**, 1463 (1982).
2. Winfree, A. T. & Strogatz, S. H. *Physica D* (submitted).

3. Winfree, A. T. *Science* **181**, 937 (1973).
4. Winfree, A. T. *JCS Faraday Trans. 9*, 38–46 (1974).
5. Winfree, A. T. *The Geometry of Biological Time*, 252–254 (Springer, New York, 1980).
6. Patony, G. & Noszticzus, Z. *React. Kinet. Catal. Lett.* **17**, 187 (1981).

Oxygen isotope ratios in German groundwater

H. Förstel & H. Hützen

Jülich, Nuclear Research Centre (KFA), Institute of Radioagronomy, D-5170 Jülich, FRG

The natural variation in the $^{18}\text{O}/^{16}\text{O}$ ratio in precipitation has a characteristic global pattern^{1,2}, a continuous decrease from the Equator towards the polar regions, corresponding to the worldwide distribution of the local mean annual temperature. This global pattern is not, however, sensitive enough to predict the local $^{18}\text{O}/^{16}\text{O}$ ratio in precipitation. A survey of groundwater throughout the FRG was obtained from about 900 samples of the municipal water supply. Using a multicorrelation study we demonstrate here that both distance from the coast and height above sea level determine the local $^{18}\text{O}/^{16}\text{O}$ ratio, and that this ratio generally decreases from the northern coast towards the southern mountainous region. The geographical variation in the $^{18}\text{O}/^{16}\text{O}$ ratio may serve as a basis for palaeoclimatological, plant-physiological and food-analytical applications. The practical applications are demonstrated by a study of the $^{18}\text{O}/^{16}\text{O}$ ratio of water in grapes and commercial wines; such observations may help to trace where a commercial wine originated.

Published results suggest a model in which subsequent precipitations are deposited as different layers within the soil^{3,4}, and are mixed during their way down through the soil column. The groundwater should therefore represent a long-term average of local precipitation, an assumption which has been confirmed for Jülich⁵.

About 900 samples from municipal water supplies of 480 stations in the FRG were analysed (supported by the German Association of Gas and Water Works, Bonn). The dense network should enable one to detect those values which do not represent the local precipitation (for example, filtrates from river banks). The samples were either taken directly from the water sources (such as wells) or from storage tanks. $^{18}\text{O}/^{16}\text{O}$ was measured by a mass spectrometer (Micromass 602 C) after an isotopic exchange between water and carbon dioxide at 293 K overnight. All results are reported as δ values (in ‰) related to Vienna-SMOW⁶:

$$\delta = \left(\frac{R_{\text{sample}}}{R_{\text{standard}}} - 1 \right) \times 10^3 \text{ ‰.}$$

where R is the $^{18}\text{O}/^{16}\text{O}$ atom ratio. The reproducibility of the method over 6 months was ± 0.2 to $\pm 0.3\text{‰}$.

The geographical distribution of the oxygen isotope ratio in groundwater (and precipitation) in Fig. 1 demonstrates the decline of the $^{18}\text{O}/^{16}\text{O}$ ratio from the northern area (coast) to the south (Bavarian Alps). Table 1 summarizes two exceptions: the River Rhine is mainly supplied by the precipitation of the southern regions of Germany. Consequently the water produced near the banks of the Rhine shows lower δ values than a sample collected in the neighbourhood near the bank of the Moselle, which represents the precipitation in the elevated region of Middle Europe. The other exception is the δ values of East Frisia, which may demonstrate the influence of seawater. In contrast to these results samples from two islands off the coast show δ values of lower ^{18}O -content; there the water was collected from precipitations.

The general trend in the FRG can be predicted sufficiently well with the aid of a multicorrelation analysis. The equation is:

$$Y = -7.1 - 2.3 \times 10^{-3} \times H - 2.8 \times 10^{-3} \times D - 1.4 \times 10^{-4} \times N \text{ ‰}$$

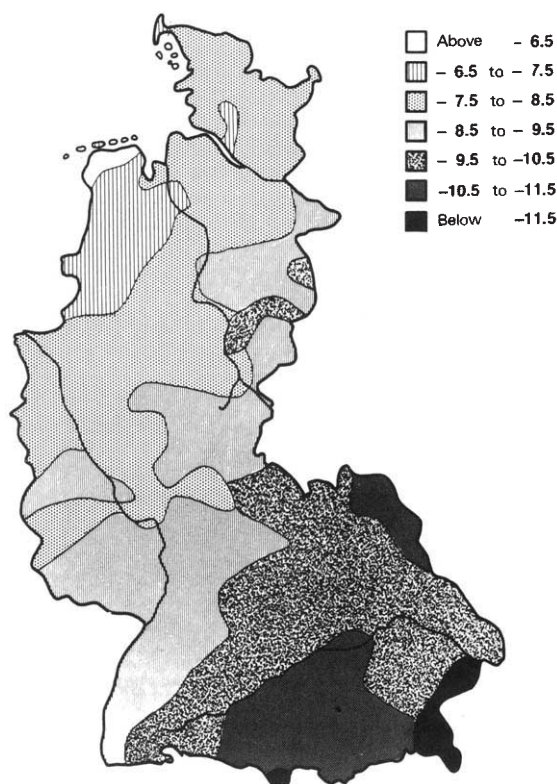


Fig. 1 Geographical distribution of the $^{18}\text{O}/^{16}\text{O}$ ratio in the FRG (isolines drawn by interpolation).

where Y is the predicted $^{18}\text{O}/^{16}\text{O}$ ratio at an observation site in ‰; H is the height above sea level in m; D is the distance from the coast in km; N is the mean annual precipitation in mm; and the correlation coefficient is 0.99.

The equation describes the groundwater (precipitation) at the coast as about -7.1‰ , which corresponds to the results from the water samples of the East Frisian Islands (see Table 1). The δ values decrease from the coast in the following manner: $\sim 2.3\text{‰}$ per km above sea level, $\sim 2.8\text{‰}$ per 1,000 km distance from the coast and 0.14‰ per m mean annual precipitation. The distance from the coast was measured parallel to the Emden-Berchtesgaden line (NW-SE).

The general pattern of the δ values is due to the subsequent fractionation steps in the water cycle: during precipitation, evapotranspiration and transport of water vapour by the air. The number of precipitation-evapotranspiration steps increases from the coast towards the land. Elevations enhance the amount of precipitation. An increase of the amount of precipitation shifts the isotopic ratio to more negative values. Our experiences with local δ values have shown that the wind direction, which is usually measured near the ground, cannot be used as an indication of the origin of air masses and their precipitation. A simple model, stepwise fractionation alongside the main direction of air flow (NW-W), has not been confirmed in our case.

The small local deviations of δ may be caused by the annual cycles of precipitation, evapotranspiration and vegetation period. In the study of the local ecology the δ value of groundwater of a station can be considered as constant. Figure 2 summarizes the results of $^{18}\text{O}/^{16}\text{O}$ ratios in Jülich during the past decade. The scattering of the δ values of precipitation diminishes on the way from a single event to a long-term mean in the groundwater. In the transport system of plants no significant fractionation can be observed. A distinct and variable fractionation of water is at least observed in the leaf⁷.

During the phase transition from fluid to vapour the heavier molecule H_2^{18}O (compared with H_2^{16}O) is enriched in the leaf

Table 1 Samples from the municipal water supply along the banks of the Lower Rhine and from the region of East Frisia

Region	Place	Source	δ (‰)
Lower Rhine	Koblenz	Bank filtrate from Rhine	-9.73
		Bank filtrate from Moselle	-7.85
	Bonn	Bank filtrate from Rhine	-9.41
		Well	-7.52
	Cologne	Bank filtrate from Rhine	-9.79
		Gravel layers	-8.46
East Frisia	Wangerooze Island	Water lens from precipitation	-7.17
	Norderney Island	Groundwater from precipitation	-7.06
	Emden	Deep well	-5.86
	Marienhaf	Deep well	-4.91
	Moorweg	Deep well	-5.50
	Norden	Groundwater	-4.50

Table 2 δ values of water in wine from different geographical origin

	Origin	δ (s.d.)	No. of observation
FRG	Ahr*	-2.7 ± 0.8	39
	Baden	-3.2 ± 1.0	14
	Württemberg	-3.2 ± 0.5	5
	Franken	-2.6 ± 1.3	5
	Mosel	-3.9 ± 0.9	38
	Nahe	-3.4 ± 0.8	12
	Rheingau	-3.7 ± 0.9	21
	Rheinhessen	-3.2 ± 0.8	60
	Rheinpfalz	-3.0 ± 1.0	70
Grapes from Rheinpfalz and Rheinhessen* (FRG)		-0.91 ± 1.22	17
USA	New York	-3.4 ± 1.2	2
	Ohio	-4.2 ± 0.6	6
	Michigan	-3.2 ± 0.6	7
	California	2.4 ± 1.3	18
Europe	Spain	2.1 ± 2.4	20
	Portugal	2.2 ± 0.3	2
	Austria	-4.3 ± 1.0	10
	Hungary	-1.5 ± 1.2	10
	France	-0.4 ± 1.6	28
	Italy	-1.0 ± 1.4	16

* Original names of German wine-growing regions.

water⁸. This enrichment, assuming steady-state conditions, is proportional to the relative humidity of the air. Long-term observations in Jülich have demonstrated that the absolute water content of the air remains constant during the day. Therefore the relative humidity of air is determined by the temperature. At least, the air temperature governs the $^{18}\text{O}/^{16}\text{O}$ ratio of leaf water. But the δ value of soil water must also be taken into account. Because it depends on the whole climatological situation of the observation station, the δ value of leaf water is determined by the whole local climate, but not by the temperature alone. The isotope value of leaf water is stored in the cellulose, and is changed by the isotope fractionation during the metabolic pathway⁹. The cellulose deposited in the tree rings can be dated very precisely.

The exchange of water vapour between leaf and air is a rapid procedure because the large surface encloses a small water volume^{7,8}. In contrast, the surface of most fruits is small compared with their water volume and the exchange is a slow procedure. Therefore the water in fruit stores an average δ value that is higher than that of the local precipitation and groundwater. Table 2 compares δ values of water from grapes and from commercial wine samples in the FRG. The difference

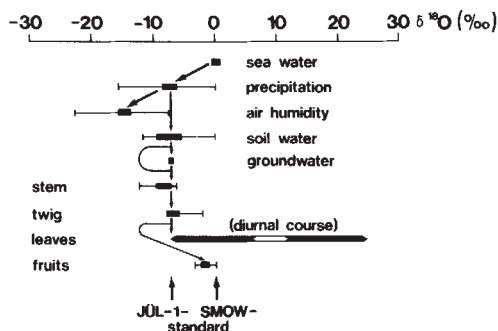


Fig. 2 $^{18}\text{O}/^{16}\text{O}$ ratios of the water in the different environmental compartments in Jülich. Solid lines characterize the common values, the light bars include extreme values. The flow of water through the local exosystem is symbolized by the arrangement of the data from top to bottom and the arrows.

between the δ value of water from grapes and from wines could result from an isotope fractionation or an exchange during fermentation. Irrespective of that difference, the wine samples from the FRG have nearly the same δ value. This is in accordance with the fact that the δ values of groundwater (and precipitation) in all wine-growing regions are within the same range.

These results are supported by measurements of wine samples from an area of similar oxygen isotope ratios in the environment¹⁰: from the wine district near the American Great Lakes (New York State, Ohio, Michigan). As a result of the different local δ value of precipitation¹, the wine samples from California have a higher $^{18}\text{O}/^{16}\text{O}$ ratio. Their δ value is comparable with results from countries of similar geographical location (Spain, Portugal).

The differences between the δ value of tap water and fruit juice should be large enough to detect the addition of water. Additionally, the geographical distribution of δ values should enable one to trace the region of origin of commercial wines.

Received 2 March; accepted 16 June 1983.

1. Yurtsever, Y. & Gat, J. R. in *Stable Isotope Hydrology* (eds Gat, J. R. & Gonfiantini, R.), 103–142 (IAEA Vienna, 1981).
2. Förstel, H., Putral, A., Schleser, G. & Lieth, H. *IAEA Symp.* **191**, 3–20 (1975).
3. Zimmermann, U., Ehalt, D. & Münnich, K. O. *IAEA Proc. Ser.* **141**, 567–585 (1967).
4. Eichler, R. *Geol. Rdsh.* **55**, 144–159 (1965).
5. Förstel, H. & Hütten, H. *Jülich Rep.* **1524** (1978).
6. Gonfiantini, R. *Nature* **271**, 534–536 (1978).
7. Dongmann, G., Förstel, H. & Wagener, K. *Nature new Biol.* **240**, 127–128 (1972).
8. Förstel, H. *Radiat. env. Biophys.* **15**, 232–344 (1978).
9. Burk, R. L. & Stuiver, M. *Science* **211**, 1417–1419 (1981).
10. Holt, B. D. *Analyt. Chem.* **49**, 1664–1667 (1977).

Ion microprobe identification of 4,100–4,200 Myr-old terrestrial zircons

**D. O. Froude, T. R. Ireland, P. D. Kinny,
I. S. Williams & W. Compston**

Research School of Earth Sciences, The Australian National University, GPO Box 4, Canberra, A.C.T. 2601, Australia

I. R. Williams & J. S. Myers

Geological Survey of Western Australia, 66 Adelaide Terrace, Perth, Western Australia 6000, Australia

We report here the existence of detrital zircons from Western Australia which are far older than any known terrestrial rocks. They are from quartzites at Mt Narryer (Fig. 1), a locality which has created interest because of the nearby occurrence of 3,630 ± 40 Myr orthogneisses¹. The older zircons were discovered during reconnaissance U–Pb age determinations of zircon concentrates from Archaean metasediments, using the ion microprobe SHRIMP^{2,3} at the Australian National Univer-

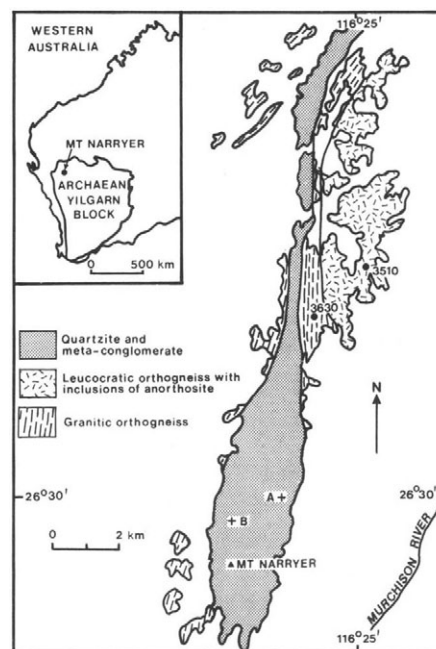


Fig. 1 Simplified geological map of the Mt Narryer region showing: ●, Location of dated gneisses with Sm–Nd model ages¹ in Myr. +, Location of quartzite samples (+A, GSWA sample site 71932; +B, GSWA sample sites 71921 and 71924).

sity. These determinations are being made specifically to search for zircons having ages in the interval 3,800–4,500 Myr, a period unrepresented so far by reliable terrestrial age determinations. Grain-by-grain measurements of one particular concentrate revealed four zircons (Fig. 2) having near-concordant U–Pb ages between 4,100 and 4,200 Myr, in striking contrast to most grains whose ages are ~3,750 and ~3,500 Myr. These results show that pre-3,800 Myr silica-saturated rocks were indeed present on the Earth's crust.

The quartzites in which the zircons occur are part of a thick sequence of waterlain clastic metasedimentary rocks. The sequence includes metaconglomerates and quartzites with well preserved cross-bedding despite metamorphism at high amphibolite to granulite facies⁴. The more argillaceous beds are now biotite-cordierite-garnet-sillimanite gneiss. The quartzite outcrop is surrounded by banded granitic gneiss (Fig. 1) which gives a Sm–Nd model age of 3,630 ± 40 Myr (ref. 1). This gneiss is intruded by another orthogneiss which contains abundant inclusions of a meta-anorthosite-gabbro-ultramafic suite and gives a Sm–Nd model age of 3,510 ± 50 Myr (ref. 1). The orthogneisses show a more complex deformation history than the quartzites, but the primary contact relations between these rocks are obliterated by local zones of intense deformation. Samples of orthogneiss define a Rb–Sr total-rock isochron of 3,350 ± 43 Myr (ref. 1) which may date regional deformation and the amphibolite-to-granulite facies metamorphism of both orthogneisses and quartzites.

The present discovery would have been impossible without the use of SHRIMP. This instrument differs from other ion microprobes mainly in its high sensitivity when operating at high mass resolution, which is achieved by the combination of wide slits and a physically-large secondary mass analyser. Its first geological applications were made in 1980⁵. For U–Pb age determinations^{2,6}, it is operated at mass-resolution 6500 to separate all significant molecular interferences from the Pb, Th and U isotopes. In addition, methods^{2,6} have been devised for the accurate measurement of Pb/U and Th/U in zircons so that $^{206}\text{Pb}/^{238}\text{U}$, $^{207}\text{Pb}/^{235}\text{U}$ and $^{208}\text{Pb}/^{232}\text{Th}$ ages can be measured, as well as the $^{207}\text{Pb}/^{206}\text{Pb}$ age. The zircon concentrates were mounted in an epoxy resin disk and polished using diamond paste until the grains were sectioned approximately in half (Fig. 3). The spatial resolution for individual analyses

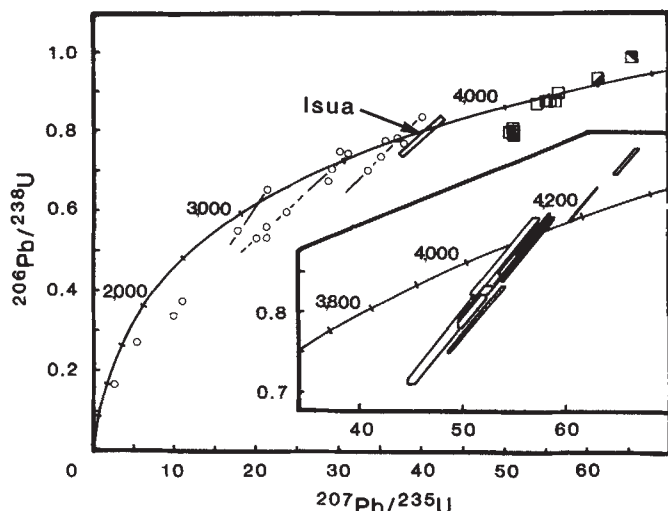


Fig. 2 Concordia diagram for zircons from quartzite 71932, Mt Narryer, Western Australia. □, Different spot analyses of grain 34 (Fig. 3); ■, ▨, ▩, the other three grains having apparent ages >4,100 Myr; ○, all other zircons. The latter appear to be of at least three different ages, as the three discordance trends show. Unpublished SHRIMP analyses of zircons from the Isua supracrustal belt are represented by the labelled rectangle. The inset gives the uncertainties in the ion microprobe results as ± 1 standard error of the spot mean. The uncertainties may include real variations in Pb/U per spot encountered during progressive sputtering, in addition to technical error.

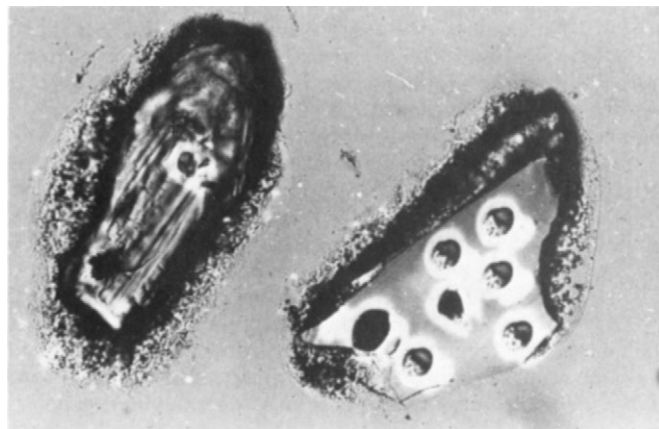


Fig. 3 Transmitted light photomicrograph of two zircons from quartzite 71932, showing grain morphology and ion microprobe sputtering pits which are $\sim 25 \mu\text{m}$ in diameter. Grain 34 (RHS) is uniquely angular, free of inclusions, and $\sim 4,110$ Myr old. Grain 35 (LHS) is typical of the morphology of the rest of the zircon population and has a $^{207}\text{Pb}/^{206}\text{Pb}$ age of 3,540 Myr. Its rounded outline is attributed to metamorphic corrosion rather than mechanical abrasion.

was $25 \mu\text{m}$ and the depth of the spot was $\sim 3 \mu\text{m}$ after a full analysis lasting ~ 45 min.

The oldest zircons occur in sample 71932 (Fig. 1). The $^{207}\text{Pb}/^{206}\text{Pb}$ model ages of two of these zircons are 4,110 Myr (Table 1); the other two are apparently older at 4,180 Myr. One of the former zircons was analysed at seven different places, one of which, 34-1, was detected as being slightly 'younger' than the others relative to experimental precision. However, using the Concordia diagram (Fig. 2), all four grains may be interpreted as having the same age, $\sim 4,150$ Myr, and having suffered redistribution of radiogenic Pb during one or more later metamorphic periods; grains 31 and 12 either gained radiogenic Pb or lost U at the analysed spots, a process that is rare in zircons but documented using SHRIMP in unpublished analyses of other zircons. We favour this interpretation because the quartzite has clearly undergone high-grade metamorphism at least as early as 3,300 Myr. In addition, some very discordant zircons (Fig. 2) show evidence of Pb loss at an intermediate age, as well as the usual 'recent' Pb loss observed in nearly all zircons.

It is conceivable that the 4,100 Myr ages might be an artefact generated by the combined effects of U loss from 3,700-Myr zircon during a later, although early, metamorphic episode (for example, at 3,300 Myr), and loss of radiogenic Pb comparatively recently. The former would displace the data above Concordia and to a higher $^{207}\text{Pb}/^{206}\text{Pb}$ model age, and the latter would displace it towards the origin. There are three points that argue against such a hypothesis. First, it has never been observed in orthodox U-Pb geochronology of zircons. Second, it is unlikely that the present data would lie, by chance, so well-grouped and close to Concordia, as observed. Third, the multiple analyses of grain 34 argue against it. There is a range of a factor of 1.5 in the U content of different areas of grain 34, yet the Pb/U and Pb isotopic composition remain virtually constant and concordant throughout. Such would be a remarkable coincidence if the Pb were, in part, unsupported.

Most of the other zircons in 71932 and in a second quartzite, 71921, formed during two younger periods, $\sim 3,750$ and $\sim 3,500$ Myr. Figure 2 contains data for the 20 zircons from 71932 that were analysed precisely. Reconnaissance scans of a further 78 zircons from 71932 for $^{207}\text{Pb}/^{206}\text{Pb}$ and $^{204}\text{Pb}/^{206}\text{Pb}$ similarly indicated ages from 3,500 to 3,750 Myr. Figure 2 also shows two younger zircons that give a near-concordant age of 3,100 Myr. Zircons from 71921 gave ages of 3,750, 3,500 and $\sim 3,300$ Myr. The $\sim 3,300$ -Myr age is within error of the age

Table 1 U-Pb analytical data for $\geq 4,100$ Myr zircons from Mt Narryer quartzite, GSWA sample 71932

Grain—spot	% f*	$^{206}\text{Pb}/^{238}\text{U}$	$^{207}\text{Pb}/^{235}\text{U}$	$^{207}\text{Pb}/^{206}\text{Pb}$	$^{207}\text{Pb}/^{206}\text{Pb}$ age (Myr)	$^{208}\text{Pb}/^{206}\text{Pb}$	U (p.p.m.)	Th (p.p.m.)
12-2	0.025	0.989 ± 18	65.7 ± 1.4	0.4819 ± 21	4,190	0.1731 ± 22	275 ± 10	195 ± 7
31-2	0.042	0.934 ± 23	61.6 ± 1.6	0.4782 ± 8	4,180	0.2829 ± 18	332 ± 9	369 ± 8
34-1	0.023	0.870 ± 49	54.1 ± 3.4	0.4507 ± 31	4,090	0.0341 ± 12	399 ± 18	56 ± 3
34-2	0.063	0.898 ± 27	56.8 ± 1.7	0.4581 ± 3	4,110	0.0389 ± 4	389 ± 12	58 ± 2
34-3	0.053	0.799 ± 36	50.8 ± 1.5	0.4605 ± 38	4,120	0.0292 ± 24	538 ± 39	65 ± 4
34-4	0.065	0.879 ± 33	55.5 ± 2.2	0.4574 ± 13	4,110	0.0280 ± 4	568 ± 14	64 ± 3
34-5	0.172	0.879 ± 41	56.3 ± 2.5	0.4644 ± 14	4,130	0.0295 ± 7	461 ± 32	58 ± 3
34-6	0.057	0.877 ± 44	55.8 ± 2.8	0.4614 ± 6	4,120	0.0263 ± 3	589 ± 12	65 ± 1
34-7	0.094	0.810 ± 22	51.2 ± 1.8	0.4583 ± 41	4,110	0.0396 ± 6	459 ± 10	67 ± 1
63-1	0.095	0.790 ± 43	51.3 ± 3.0	0.4703 ± 23	4,150	0.2039 ± 23	228 ± 8	174 ± 9
63-2	0.385	0.809 ± 99	51.2 ± 6.7	0.4581 ± 60	4,110	0.1469 ± 28	564 ± 61	312 ± 37

The uncertainty shown is $\pm 1\sigma$ and refers to the last two digits; it may include real dispersion in the target during analysis, as well as experimental error.

* Mean value of initial $^{206}\text{Pb}/\text{total } ^{206}\text{Pb}$ for 21 data cycles, based on the observed ^{204}Pb and assuming mean terrestrial Pb at 4,100 Myr.

of regional metamorphism inferred from Rb–Sr analyses of the nearby orthogneisses. We therefore view both the 3,100 and 3,300 Myr zircons as having grown during later metamorphism, and infer that the sedimentation is earlier than 3,300 Myr. If the older zircons are regarded as detrital, a maximum limit to the age of sedimentation is set by the 3,500-Myr group. None of the zircons show the small-scale surface chipping that is characteristic of mechanical abrasion. Most of the crystals are rounded, but this is probably a metamorphic effect. Textural study and further ion microprobe analyses of zircons *in situ* in the quartzite will be necessary to confirm these interpretations.

The veracity of the ion microprobe for measurement of $^{207}\text{Pb}/^{206}\text{Pb}$ and Pb/U is demonstrated by our cited results for the Isua acid volcanics (Fig. 2), which show identical ages and the same small range of Pb loss as the new results of Baadsgaard *et al.*⁷, who used mass-spectrometric isotopic dilution on multi-grain samples. Table 1 shows that the measured Th/U in the various quartzite zircon grains is highly correlated with the radiogenic $^{208}\text{Pb}/^{206}\text{Pb}$, demonstrating both the accuracy of the

ion microprobe determinations of Th/U and the presence of large differences in Th/U both between and within single grains. However, there are also second-order effects in the Th–Pb systematics which indicate later redistribution of radiogenic Pb and/or Th. These effects, together with $^{208}\text{Pb}/^{232}\text{Th}$ ages, will be discussed in more detail elsewhere.

The present results demonstrate the former existence of Archaean rocks at least as old as 4,100 Myr in the vicinity of Mt Narryer. It is possible that intact remnants of these rocks may yet be found.

The suitability of Mt Narryer quartzites for this study was suggested by A. F. Trendall and D. Gee; their invitation to visit Western Australia and their participation in field work is gratefully acknowledged. I.R.W. and J.S.M. publish with the permission of the director of the Geological Survey of Western Australia. D.O.F., T.R.I. and P.D.K. hold scholarships from the Commonwealth Fellowship and Scholarship Plan, The Australian National University, and the John Conrad Jaeger bequest, respectively.

Received 18 May; accepted 6 July 1983.

1. DeLaeter, J. R. *et al. Nature* **292**, 322–324 (1981).

2. Clement, S. W. J., Compston, W. & Newstead, G. *Proc. int. S.I.M.S. Conf.*, Meunster (ed. Benninghoven, A.) (Springer, Berlin, 1977).

3. Compston, W., Williams, I. S. & Clement, S. W. J. *Proc. 30th Am. Soc. Mass Spectrom. Conf.* Honolulu, 593–595 (Am. Soc. Mass Spectrom., 1982).

4. Williams, I. R., Walker, I. W., Hocking, R. M. & Williams, S. J. *Geol. Surv. Western Australia Rec.* **5** (1980).

5. Coles, J. W. N., Compston, W. & Foster, J. J. *Research School of Earth Sciences, A.N.U., A. Rep.*, 175–179 (1980).

6. Compston, W., Williams, I. S. & Meyer, C. *Proc. 14th Lunar Sci. Conf.* (in the press).

7. Baadsgaard, H. *et al. Earth planet. Sci. Lett.* (in the press).

Inter-layered clay stacks in Jurassic shales

K. Pye

Department of Earth Sciences, Downing Street,
Cambridge CB2 3EQ, UK

D. H. Krinsley

Department of Geology, Arizona State University, Tempe,
Arizona 85287, USA

As the average shale contains ~60% clay minerals¹, an understanding of the physical and chemical characteristics of these minerals is central to the question of burial diagenetic reactions and hydrocarbon generation. Much existing evidence concerning burial diagenesis relies on X-ray diffraction (XRD) data, particularly of the clay-size (<2 µm) fraction of shales². Here we use scanning electron microscopy (SEM) in the backscattered electron (BSE) mode, together with energy-dispersive X-ray microanalysis (EDXA), to show that Lower Jurassic shales from the North Sea Basin contain large numbers of clay mineral stacks up to 150 µm in size. By studying polished shale sections with BSE and EDXA, the size, shape orientation, textural relationships and internal compositional variations of the clays can be observed *in situ*. We present here preliminary evidence which suggests that the clay stacks studied are authigenic, and may have formed at shallow burial depths during early diagenesis.

Clay minerals in shales have been studied for many years using XRD and, more recently, transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM). XRD is highly important for identifying and quantifying the structural types of clay minerals present, but this technique provides little information about crystal size, shape, textural relationships or chemical composition. Consequently the morphological and textural properties of clays in shales have been relatively neglected. As XRD is conventionally carried out on the <2-µm fraction after the rock has been disaggregated by ultrasonic or other methods, this procedure may overlook clay minerals which are present as silt or sand-size grains. Attempts to explain the mineral transformations which occur during diagenesis have commonly ignored the possibility that clay minerals might increase significantly in size.

TEM and STEM can provide useful information about the size, shape and textural relationships of small crystals in mudrocks, but the area of specimen which can be examined is extremely small and an overall impression of the rock micro-fabric is difficult to obtain. It is often difficult to relate TEM observations to those made at higher magnifications with the optical microscope. SEM examination of polished mudrock sections using backscattered electrons provides a means of bridging the gap between optical microscopy and TEM. The emission of backscattered electrons from a polished surface is strongly dependent on the average atomic number of the target, such that different minerals may display contrasting grey levels in the BSE SEM image^{3–6}. This technique has three principal advantages compared with optical microscopy: (1) higher resolution; (2) subtle variations in chemical composition can be detected with BSE which would be missed with the optical microscope; and (3) the nature of the chemical variations can be determined directly using an X-ray analyser attachment on the SEM.

We have recently examined several polished sections of Lower Jurassic (Lias) shales from the Yorkshire coast and

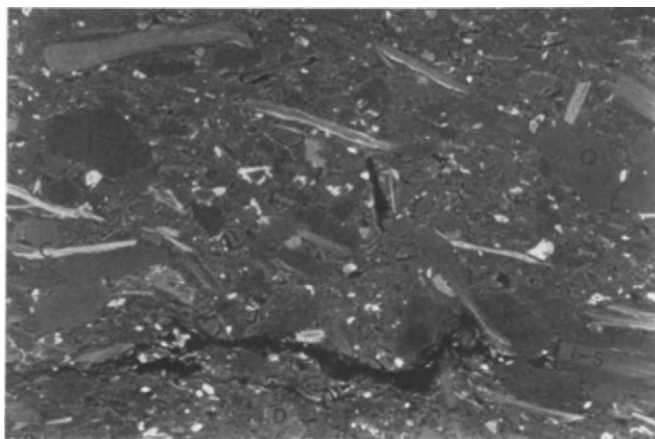


Fig. 1 BSE micrograph of a polished section of Lias shale showing clay mineral stacks orientated sub-parallel to bedding. Several types of clay are present including chlorite (C), illite-smectite (I-S) and kaolinite (K). The finer-grained matrix is also a mixture of these minerals. Also visible in this picture are quartz (Q), albite (A), and zoned carbonate crystals (D).

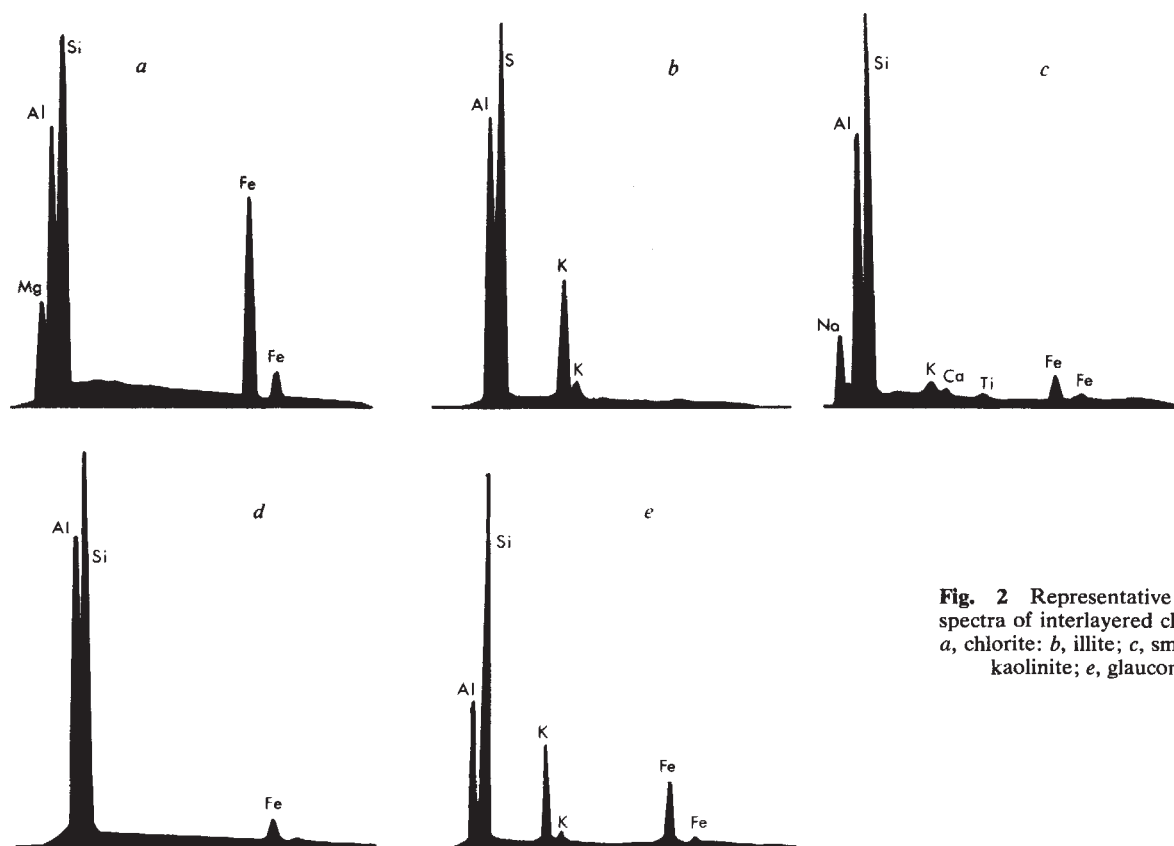


Fig. 2 Representative EDXA spectra of interlayered clay types: *a*, chlorite; *b*, illite; *c*, smectite; *d*, kaolinite; *e*, glauconite.

offshore sector of the southern North Sea using BSE and EDXA in SEM as part of a study of diagenesis in Mesozoic mudrocks. Full results of this investigation will be published elsewhere. We point out here that the Lias samples examined contain numerous clay mineral stacks which range in size up to 150 μm , and demonstrate that BSE-EDXA techniques are a useful aid in interpreting the origin and nature of interlayering in such minerals.

The Lias shales examined are predominantly fine-grained mudstones, in places organic-rich, which have been buried to depths of between 2.5 (ref. 7) and 3.2 km. A range of techniques has been used to characterize the material, including X-ray fluorescence (XRF), XRD, microprobe, optical microscopy, TEM and SEM-EDXA. The major minerals present are illite, smectite, chlorite, kaolinite (including interlayered combinations), quartz, feldspars, pyrite, apatite and carbonates. All sizes of clay mineral are present from <0.01 μm to large blocky grains more than 100 μm long and 40 μm wide. Figure 1 illustrates a typical association of large, elongated clay mineral stacks within a matrix of finer clay. In the BSE image iron-rich varieties of chlorite normally appear light grey or white, illite is mid-grey, and kaolinite and smectite are relatively dark. As can be seen in Fig. 1, many of the clay stacks show compositional interlayering. Using EDXA, supported by XRD and electron diffraction data, 10 different clay mineral associations have been identified: illite, chlorite, smectite, kaolinite, glauconite, illite-smectite, chlorite-smectite, chlorite-illite, illite-kaolinite and illite-smectite-chlorite. Of these, the most common interlayered clay types are illite-smectite, chlorite-smectite and illite-chlorite. Representative EDXA spectra of the main types of clay inter-layer are illustrated in Fig. 2.

Viewed in sections cut transverse to the bedding, the shape of the clay stacks ranges from elongate blades or sheets (length-to-width ratios up to 20:1) to equidimensional blocky grains (length-to-width ratios of 1.5:1 to 3:1). Although there appears to be no universal rule, many chlorite grains take the form of long, thin sheets, illite-smectite grains often occur as long but thicker stacks, and kaolinite grains are frequently short

and blocky. In some cases the thickness of the stacks varies significantly along the length of the crystal, and a few examples bifurcate to give a tuning fork or feathered appearance. Glauconite invariably occurs as coarse silt or sand-size rounded pellets.

In some stacks the interlayering is regular and broadly laminar, but in many other cases it is wavy and exhibits bifurcation (Fig. 3). Several grains have been observed with clear discontinuities on either side of which the interstratification is inclined at a different angle. In other examples a central core of illite, exhibiting no evident interlayering, is surrounded by a rim of 'pure' chlorite. EDXA has shown that not all of the colour bands represent an entirely different mineral species, but rather correspond to slight variations in the concentrations of cations present. Examples have been noted of alternating Mg and Fe-rich chlorite, K-rich and K-poor illite, and Na-rich and Na-poor smectite.

The question of the diagenetic transformations in these rocks will be considered more fully elsewhere, but a brief discussion of the possible origin of the stacks is pertinent. Three main possibilities exist: first, the stacks might be detrital; second, they might represent detrital micas which have been altered post-depositionally; third, they may be authigenic. Several lines of evidence suggest that a detrital origin is unlikely. First, the very large size of many of the stacks compared with that of the surrounding quartz and feldspar grains makes it improbable that they were deposited in equilibrium with hydraulic conditions, even bearing in mind that clay stacks have lower specific gravities than quartz and should be expected to be coarser-grained as hydraulic equivalents. In some instances the clay stacks are more than 100 times larger than the surrounding quartz grains. Second, many stacks display delicate structures and projections which are too fragile to have survived long-distance transport (Fig. 3). The unusual shapes and complex interlayered nature of certain stacks suggest they could not have been formed by differential alteration of detrital micas. Third, particularly in the organic-rich facies, many clay stacks contain inclusions of pyrite, quartz and other minerals which



Fig. 3 BSE micrograph showing complex nature of interlayering in a chlorite-smectite stack. The delicate growth projection at the right-hand end of the stack is unlikely to have survived long-distance transport.

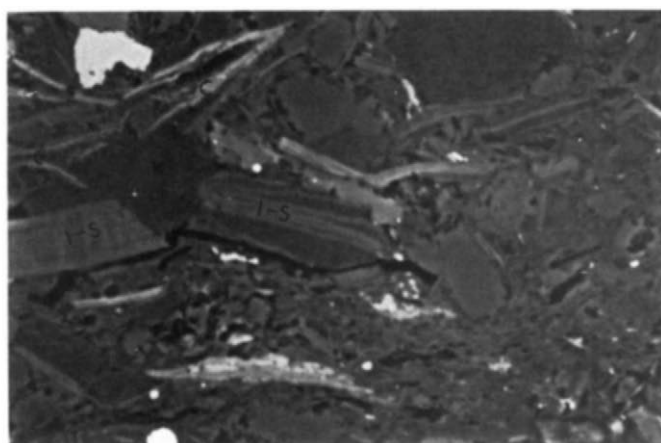


Fig. 4 BSE micrograph showing authigenic intergrowth of illite-smectite (I-S), kaolinite (K) and chlorite (C).

appear to have been incorporated during growth. Finally, groups of chlorite, illite and kaolinite crystals sometimes show complex textural intergrowths with each other and with quartz, feldspars, and authigenic dolomite-ankerite crystals (Fig. 4). We feel that the most likely explanation of these textural intergrowths is authigenic development of the clay stacks. The relatively large size and apparent fragility of the aggregates makes a detrital origin unlikely, and it is difficult to see how such features could be produced by differential alteration of detrital micas.

Further evidence suggests that growth of the clay stacks may have occurred at shallow depths during early diagenesis. First, the long axes of the stacks show a general sub-parallelism to the bedding (Fig. 1), except in the vicinity of what appear to be small burrows where the sediment fabric has been disturbed. Second, the finer-grained matrix surrounding the larger stacks often shows signs of deformation, suggesting growth of the stacks in soft, uncompacted sediment. Third, the concentration of clay stacks is often highest just below organic-rich laminae, which probably represent periods when inorganic sedimentation slowed or ceased. Finally, in several cases clay stacks have been observed to be cemented by glauconite or pyrite which are probably of near-surface origin. Clay stacks have also been observed to occur within calcite concretions, features which are generally accepted on the basis of isotopic and other evidence to be early diagenetic^{8,9}. Although we consider that the hypothesis is not yet fully proven, it seems possible that most clay mineral stacks formed at, or very close to, the sediment-water interface during periods of slow sedimentation. Diagenetic reactions are generally considered to be of minor importance in explaining the pattern of clay mineral assemblages in the modern oceans¹⁰⁻¹³, although some authors have postulated large-scale neo-formation of clay minerals pencon-temporaneous with deposition¹⁴. Early authigenic chamosite and glauconite are found in a variety of modern marine environments¹⁵⁻¹⁸, and there appears to be no reason why other clay mineral stacks should not form in the right geochemical conditions.

We have attempted to show that BSE-EDXA analysis of shale thin sections can provide useful information about the nature of interstratification in clays as well as textural data which help to elucidate their origin. The SEM provides a valuable link between the optical microscope and bulk chemical-mineralogical data on the one hand and TEM results on the other. Our observation that clay minerals may exist in Jurassic shales as silt or fine sand-size particles suggests that conventional XRD procedures need to be modified to obtain a fully representative picture. A tremendous variation exists in size, shape and internal structures of clay minerals, as well as other fine-grained constituents of shales, which has not formerly

been fully recognized. Although this creates problems for the mudrock petrographer, it also offers hope that very detailed physical and chemical changes with time can be observed and understood.

K.P. acknowledges financial support provided by a NERC Research Fellowship and Girtton College. D.H.K. was supported partly by NASA-Ames research grant NAG 2-174 and the Planetary Geology Program, Office of Space Science, research grant NAGW24. We thank David Newling, Pat Trusty, Bill Porter and Brian Krinsley for technical assistance, and British Petroleum for provision of some offshore samples.

Received 25 April; accepted 15 June 1983.

1. Shaw, D. B. & Weaver, C. E. *J. sedim. Petrol.* **35**, 213-222 (1965).
2. Hower, J., Eslinger, E. V., Hower, M. E. & Perry, E. A. *Bull. geol. Soc. Am.* **87**, 725-737 (1976).
3. Hall, M. G. & Lloyd, G. E. *Am. Miner.* **66**, 362-368 (1981).
4. Pye, K. & Krinsley, D. H. *Nature* **301**, 412-413 (1983).
5. Krinsley, D. H., Pye, K. & Kearsley, A. T. *Geol. Mag.* **120**, 109-114 (1983).
6. Pye, K. & Krinsley, D. H. *J. sedim. Petrol.* (in the press).
7. Hemingway, J. E. & Riddler, G. P. *Trans. Instn. Min. Metall. B. appl. Earth Sci.* **91**, B175-B186 (1982).
8. Raiswell, R. *Chem. Geol.* **18**, 227-244 (1976).
9. Coleman, M. L. & Raiswell, R. *Geochim. cosmochim. Acta* **45**, 329-340 (1981).
10. Weaver, C. E. *Bull. Am. Ass. petrol. Geol.* **42**, 254-271 (1958).
11. Biscaye, P. E. *Bull. geol. Soc. Am.* **76**, 803-832 (1965).
12. Griffin, J. J., Windom, H. & Goldberg, E. D. *Deep Sea Res.* **15**, 433-459 (1968).
13. Rateev, M. A., Gorbunova, Z. N., Lisitzyn, A. P. & Nosov, G. L. *Sedimentology* **13**, 21-43 (1969).
14. Jeans, C. V. *Clay Min.* **9**, 209-217 (1971).
15. Porrenga, D. H. *Mar. Geol.* **5**, 495-501 (1967).
16. Rohrlach, V., Price, N. B. & Calvert, S. E. *J. sedim. Petrol.* **39**, 624-631 (1969).
17. Hein, J. R., Ailwardt, A. O. & Griggs, G. B. *J. sedim. Petrol.* **44**, 562-571 (1974).
18. Birch, G. F., Willis, J. P. & Richard, R. S. *Mar. Geol.* **22**, 271-284 (1976).

Compression, nonstoichiometry and bulk viscosity of wüstite

Raymond Jeanloz

Department of Geology and Geophysics, University of California, Berkeley, California 94720, USA

Robert M. Hazen

Geophysical Laboratory, Carnegie Institution of Washington, Washington DC 20008, USA

Wüstite (Fe_{1-x}O) is a complex oxide that is invariably non-stoichiometric ($0.04 \geq x \geq 0.12$). Point defects associated with this nonstoichiometry are clustered and result in a modulated crystal structure that has been extensively studied, but is not well understood¹⁻⁴. Many of the properties of these compounds, such as chemical diffusivity, electrical conductivity and magnetic transitions, seem to be significantly affected by the non-

stoichiometry and resulting ordering of defects⁵⁻⁸. We summarize here new and existing data on the elasticity of wüstite; these data indicate virtually no correlation between the bulk modulus and stoichiometry. There appears, however, to be a systematic difference between static and dynamic measurements. This behaviour suggests the presence of a frequency dependence to the bulk modulus that can be ascribed formally to a finite bulk viscosity of Fe_{1-x}O . Such an effect has not previously been described in oxide or silicate minerals, and it is presumably caused by pressure-induced changes in defect ordering and subsequent relaxation within the incommensurate structure of wüstite.

The bulk modulus of wüstite has been determined by several techniques, as illustrated in Fig. 1. Dynamic measurements include ultrasonic and shock-wave experiments that probe the sample on time scales of microseconds or less. In contrast, static determinations are carried out over periods of 10^5 – 10^6 s. Jackson *et al.*⁹ extrapolated ultrasonic data on polycrystalline (Mg, Fe)O samples to yield a bulk modulus of FeO that is in agreement with Mizutani's earlier determination on polycrystalline wüstite^{9,10} and the more recent single-crystal measurement of Sumino *et al.*¹¹ (based on cube-resonance interferometry). These results are consistent with the shock-compression data of Jeanloz and Ahrens¹² on polycrystalline wüstite in defining a dynamic value for the zero-pressure isothermal bulk modulus of wüstite:

$$K_{OT}^{\text{Dynamic}} = 178(\pm 6) \text{ GPa}$$

The hydrostatic compression of wüstite has recently been studied in two independent experiments^{13,14}. The measurements are in agreement with previous, nonhydrostatic determinations of the bulk modulus of wüstite¹⁵⁻¹⁷. Thus, we find a value of the static bulk modulus:

$$K_{OT}^{\text{Static}} = 155(\pm 5) \text{ GPa}$$

Nonhydrostatic conditions, which obtain in some of these experiments, do not seem to affect the results, contrary to what might be expected from previous static-compression studies¹⁸. Note that three pressure calibrations were utilized: Mao *et al.*¹⁷ and Will *et al.*¹⁶ used NaCl and Fe, respectively, as pressure standards, whereas the ruby-fluorescence technique was employed in the hydrostatic experiments¹³⁻¹⁷. This multiplicity of pressure standards reduces the possibility of systematic bias among the static-compression studies. We have used Eulerian finite-strain theory (Birch equation of state) to reduce the data of Will *et al.* and Mao *et al.* in accord with the analysis of the hydrostatic-compression data^{14,19}. Only the low-pressure data of Mao *et al.* have been considered because of the evidence for distortion of the crystal structure of wüstite that is associated with a second-order transition at high pressures (at or above about 10–20 GPa, depending on the degree of uniaxial stress)^{14,20}.

There is no compelling evidence among either the dynamic or static data for any change in bulk modulus with stoichiometry. In particular, we note the results of Hazen's experiment that was carried out simultaneously on three samples ($x = 0.055$, 0.07 and 0.10)¹³. The relative values of the compression of these samples (which are more precisely determined than their absolute values) yield $dK_{OT}/dx = 64 \text{ GPa}$, as shown by the dot-dash curve in Fig. 1. This slope is in agreement with several model calculations¹⁴ that lead to predictions of a value for $dK_{OT}/dx \sim \pm 0.5 K_{OT}$, as illustrated by the shaded region in Fig. 1. Continuum models, on the one hand, lead to predictions of an increase in K_{OT} with increasing x , because the volume of wüstite decreases with increasing nonstoichiometry^{11,21,22}. Atomistic models^{23,24}, on the other hand, lead to predictions of a slight decrease of the bulk modulus as x increases, as does a model that treats the defect clusters in wüstite as magnetite-like microdomains¹⁴. The relative change in bulk modulus can be as large as $\pm 100\%$ of the change in x (ref. 14); the sign and

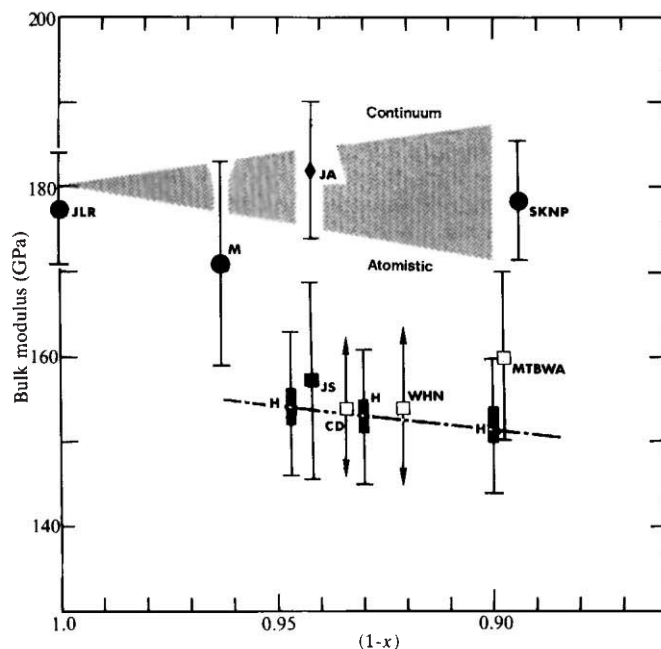


Fig. 1 Isothermal bulk modulus of wüstite as a function of composition or nonstoichiometry. Dynamic measurements are indicated by circles and a diamond, whereas static measurements are shown as squares and rectangles. Composition was determined from the lattice parameter quoted in each study^{29,30}. The techniques used include ultrasonic and cube-resonance interferometry (circles: JLR, Jackson *et al.*⁹; M, Mizutani¹⁰; SKNP, Sumino *et al.*¹¹), shock-wave compression (diamond: JA, Jeanloz and Ahrens¹²), hydrostatic compression (filled square and rectangles: H, Hazen¹³; JS, Jeanloz and Sato-Sorensen¹⁴) and nonhydrostatic compression (open squares: CD, Clendenen and Drickamer¹⁵; WHN, Will *et al.*¹⁶; MTBWA, Mao *et al.*¹⁷). The shaded region illustrates the effect of nonstoichiometry on the bulk modulus that is expected according to several theoretical and empirical models discussed in the text. (A value of 180 GPa at $x = 0.0$ is assumed.¹¹) The dependency of the bulk modulus on x observed by Hazen¹³ is shown by the dotted-dashed line (relative uncertainties among these determinations of bulk moduli are indicated by the sizes of the rectangles). Adiabatic moduli have been corrected to isothermal values, and static-compression measurements are all reduced according to finite-strain theory. Also, the data of Mao *et al.*¹⁷ have been reanalysed, as discussed in the text.

magnitude of the calculated change depend on the choice of parameters used in these calculations. In the light of data shown in Fig. 1, however, it appears that the bulk modulus of wüstite is essentially independent of stoichiometry. It is not possible, therefore, to ascribe the systematic difference between static and dynamic values of the bulk modulus to a dependency of K_{OT} on sample composition, x . This conclusion is supported by the differing results of the experiments of Jeanloz and Ahrens¹², and Jeanloz and Sato-Sorensen¹⁴, that were carried out on the same sample material.

The difference between static and dynamic moduli can be interpreted as a modulus defect, $\Delta K \sim 23 \text{ GPa}$ (ref. 25). If this interpretation is correct, then the data summarized in Fig. 1 imply that wüstite exhibits a bulk viscosity of magnitude $\kappa \sim \Delta K \tau$, where τ is the characteristic time for the relaxation mechanism that is involved. We cannot distinguish between anelastic and viscoelastic types of relaxation from these data, so we refer to bulk viscosity in only a loose sense without necessarily implying that finite volumetric strains of any significant magnitude need occur. As structural relaxation is observed in experiments lasting days to weeks, the magnitude of τ would be expected to be about 10^5 – 10^6 s, or less. In fact, a value $\tau \sim 10^6$ s is consistent with the data of Hazen *et al.*, which suggest that wüstite undergoes an irreversible (thus,

viscoelastic) change in lattice parameter at high pressures²⁶. A rough estimate of τ can be made from high-temperature diffusivity measurements, which may be used to determine the atomic jump frequency Γ . An extrapolation of the data of Chen and Peterson⁶ from $\sim 1,100$ K (activation energy $Q = 123 \text{ kJ mol}^{-1}$) combined with the assumption that the diffusive length scale is the shortest distance between sites in the defect clusters ($\sim 0.03 \text{ nm}$) yields a predicted value of $\tau = \Gamma^{-1} \sim 10^7 \text{ s}$ at room temperature. (We use Morin's²⁷ value of ~ 0.6 for the correlation factor.) Alternatively, τ is $\sim 10^6 \text{ s}$ if the activation energy of defect migration $Q \sim 115 \text{ kJ mol}^{-1}$, which is within the range that has been reported^{5,6,28}. Therefore, we can estimate the magnitude of the bulk viscosity in wüstite $\kappa \sim 10^{16} - 10^{17} \text{ Pa s}$. Similar effects might be observed in the magnesio-wüstite series (Mg, Fe)O, because samples with as much as 40% Mg substitution for Fe display nonstoichiometric characteristics^{9,29}.

The unusual occurrence of bulk viscosity in a simple oxide is consistent with the complex defect structure that is found in wüstite^{2,4}. We infer that defect clusters are rearranged under hydrostatic compression, such that the bulk viscosity is a macroscopic description of what is, on a microscopic level, a pressure-induced, order-disorder transformation of the defect structure. This phenomenon is not a polymorphic transition; the main NaCl-type crystal structure of wüstite remains unchanged to pressures of $\sim 10 \text{ GPa}$ or more^{14,20,26}. Consequently, we would expect no difference between static and dynamic moduli for a hypothetical defect-free FeO. (Note that defects can occur in stoichiometric wüstite if equal numbers of vacancies and interstitials are present; defect relaxation may thus be observed even for $x = 0$.) We therefore consider the dynamic modulus, rather than the static modulus, to be representative of interatomic bonding forces in wüstite. The observed modulus defect provides a new constraint on models of the defect-ordering behaviour within this nonstoichiometric compound.

We thank D. Heinz, M. S. T. Bukowski, L. W. Finger, H. K. Mao and H. S. Yoder Jr. for helpful comments. This work was supported by NASA grant NAGW-439, NSF grant EAR 81-15517 and the A. P. Sloan Foundation.

Received 25 April; accepted 15 June 1983.

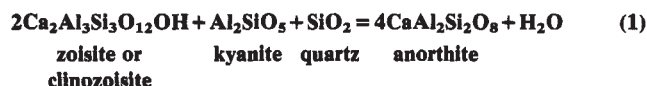
- Koch, F. & Cohen, J. B. *Acta crystallogr.* B25, 275-287 (1969).
- Andersson, B. & Sletnes, J. O. *Acta crystallogr.* A33, 268-276 (1977).
- Gavarrí, J., Carel, C. & Weigel, D. J. *Solid State Chem.* 29, 81-95 (1979).
- Yamamoto, A. *Acta crystallogr.* B38, 1451-1456 (1982).
- Kofstad, P. *Nonstoichiometry, Diffusion and Electrical Conductivity in Binary Metal Oxides* (Wiley, New York, 1972).
- Chen, W. K. & Peterson, N. L. *J. phys. Chem. Solids* 36, 1097-1103 (1975).
- Koch, F. B. & Fine, M. F. J. *appl. Phys.* 38, 1470-1471 (1966).
- Greenwood, N. N. in *Perspectives in Mössbauer Spectroscopy* (eds Cohen, S. G. & Pasternak, M.) 27-39 (Plenum, New York, 1973).
- Jackson, I., Lieberman, R. C. & Ringwood, A. E. *Phys. Chem. Miner.* 3, 11-31 (1978).
- Mizutani, H., Hamano, Y. & Akimoto, S. *J. geophys. Res.* 77, 3744-3749 (1972).
- Sumino, Y., Kumazawa, M., Nishizawa, O. & Pluschke, W. J. *Phys. Earth* 28, 475-495 (1980).
- Jeanloz, R. & Ahrens, T. J. *Geophys. J. R. astr. Soc.* 62, 505-528 (1980).
- Hazen, R. M. *Yb. Carnegie Instn Wash.* 80, 277-280 (1981).
- Jeanloz, R. & Sato-Sorensen, Y. *EOS* 63, 1097 (1982).
- Clendenen, R. L. & Drickamer, H. G. *J. chem. Phys.* 44, 4223-4228 (1966).
- Will, G., Hinze, E. & Nuding, W. *Phys. Chem. Miner.* 6, 157-167 (1980).
- Mao, H. K., Takahashi, T., Bassett, W., Weaver, J. S. & Akimoto, S. *J. geophys. Res.* 74, 1061-1069 (1969).
- Wilburn, D. R. & Bassett, W. A. *Am. Miner.* 63, 591-596 (1978).
- Birch, F. *Phys. Rev.* 71, 809-824 (1947).
- Zou, G. T., Mao, H. K., Bell, P. M. & Virgo, D. *Yb. Carnegie Instn Wash.* 79, 374-376 (1980).
- Holder, J. & Granato, A. V. *Phys. Rev.* 182, 729-741 (1969).
- Holder, J. & Granato, A. V. in *Physical Acoustics* Vol. 8 (eds Mason, W. P. & Thurston, R. N.) 237-277 (Academic, New York, 1971).
- Ludwig, W. *Natn. Bur. Stand. Misc. Publ.* 287, 151-156 (1966).
- Pistorius, M. Z. *angew. Phys.* 29, 145-152 (1970).
- Nowick, A. S. & Berry, B. S. *Anelastic Relaxation in Crystalline Solids* (Academic, New York, 1972).
- Hazen, R. M., Mao, H. K., Finger, L. W. & Bell, P. M. *Yb. Carnegie Instn Wash.* 80, 274-277 (1981).
- Morin, F. *Scripta metall.* 10, 553-556 (1976).
- Chen, W. K. & Peterson, N. L. *J. Phys. Chem. Solids* 41, 335-339 (1980).
- Simons, B. *Yb. Carnegie Instn Wash.* 79, 376-380 (1980).
- Hentschel, B. Z. *Naturf.* 25a, 1996-1997 (1970).

Fe-free clinozoisite stability relative to zoisite

D. M. Jenkins, R. C. Newton & J. R. Goldsmith

Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA

Orthorhombic zoisite, $\text{Ca}_2\text{Al}_3\text{Si}_3\text{O}_{12}\text{OH}$, and its monoclinic polymorph, clinozoisite, are key indicator minerals in low- and middle-grade basic and calcareous metamorphic rocks, and often occur together in apparent equilibrium¹. Both minerals contain variable amounts of Fe^{3+} replacing Al, but clinozoisite always contains more Fe^{3+} , with as much as 30-40% substitution in typical epidote. Interpretations of their relative stabilities, based on field relations¹⁻³ and experimental studies^{4,5}, have conflicted markedly. No unambiguous experimental demonstrations of reversibility of a zoisite-clinozoisite equilibrium have yet been performed, and it may well be that such an equilibrium relation, if it exists, lies in a portion of the P, T plane which is inaccessible to direct investigation. We present here experimental reversals of the Fe-free equilibrium:



using both clinozoisite and zoisite. At 550°C , the reaction with zoisite is in equilibrium at $6.30 \pm 0.25 \text{ kbar}$ and the reaction with clinozoisite is in equilibrium at $7.30 \pm 0.15 \text{ kbar}$. Therefore, clinozoisite is considerably less stable than zoisite at this temperature, and the reaction with clinozoisite, though reversible, is metastable. Reliable thermodynamic data for the phases show, with the present experimental data, that zoisite has lower Gibbs free energy at 550°C by $1,020 \pm 400 \text{ cal mol}^{-1}$. Published composition data of natural coexisting zoisite and clinozoisite, together with first-order assumptions as to the equilibrium temperatures of the parageneses, suggest that zoisite entropy is $\sim 2.5 \text{ cal mol}^{-1} \text{ K}^{-1}$ higher than that of clinozoisite and that Fe-free clinozoisite becomes stable relative to zoisite below about 200°C .

An internally-heated gas-pressure vessel with Ar medium was used to investigate reaction (1). Pressures were measured accurately to $\pm 50 \text{ bar}$ with both a manganin pressure cell and a bourdon-tube gauge. Two Cr-Al thermocouples at the ends of the sample showed $1-2^\circ\text{C}$ differences. Run durations were 4 days.

Starting materials were, with the exception of clinozoisite and quartz, crystalline phases synthesized from reagent-grade oxides or hydroxides. Natural quartz was from Lisbon, Maryland. Clinozoisite plus quartz was synthesized from a natural scolecite, $\text{CaAl}_2\text{Si}_3\text{O}_{10} \cdot 3\text{H}_2\text{O}$ at 500°C and 15 kbar for 4 h. The scolecite contained $<0.01 \text{ wt\% Na}_2\text{O}$ as the principal impurity. Two homogenized reversal mix powders were made which consisted of nearly equal proportions of reactants and products in reaction (1), one mix with zoisite and the other with clinozoisite. Portions of each mix were sealed in Pt tubes with $\sim 25\% \text{ H}_2\text{O}$ and placed side-by-side in the pressure vessel. The direction of a reaction produced during a run was deduced from changes of X-ray diffraction peak heights.

The experimental results at 550°C are shown in Fig. 1. Only diffraction patterns showing large apparent changes, $\sim 30\%$ in the relative amounts of product and reactant phases, are shown as defining the directions of reaction. Between 6.6 and 7.2 kbar , zoisite grew relative to anorthite, but clinozoisite broke down. Therefore, zoisite is stable with kyanite and quartz to lower pressures than clinozoisite by at least 600 bar and at most $1,400 \text{ bar}$ at 550°C .

Run charges with the clinozoisite mix were checked for the spontaneous appearance of zoisite by searching for the distinc-

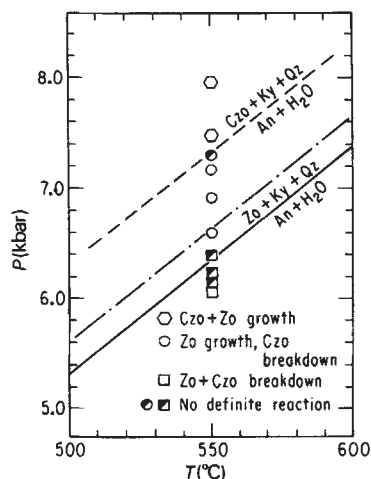


Fig. 1 Experimental reversal runs on the reaction of zoisite + kyanite + quartz (solid line) to anorthite and H_2O , and of clinozoisite + kyanite + quartz (dashed line) to anorthite + H_2O . Dot-dash line is experimental curve for the reaction with zoisite from ref. 7. Czo, clinozoisite; Zo, zoisite; Ky, kyanite; Qz, quartz; An, anorthite.

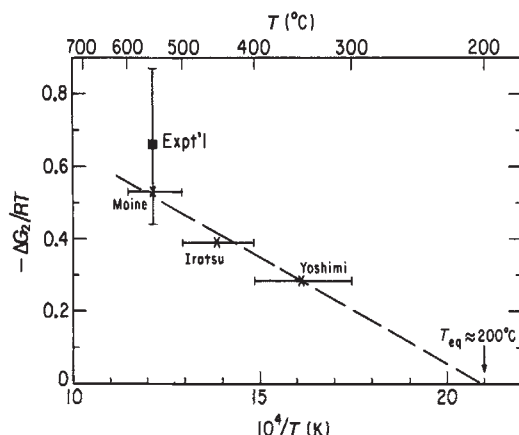


Fig. 2 Equilibrium constants of the reaction Fe-free clinozoisite = zoisite inferred from natural parageneses (see text), compared with present experimental value.

tive 013 and 213 X-ray peaks⁹. No zoisite appeared in the runs at 550 °C. However, zoisite appeared in some runs made at 600 °C, indicating that the spontaneous nucleation time of zoisite at that temperature is too short to assure metastable reversals. The uncertainty intervals of 300–500 bar at 550 °C suggest that much longer experimental times would be necessary to produce definitive reactions near the boundaries at 500 °C. Therefore, the dP/dT slopes of the reactions are not established experimentally. However, one would predict a 2.4 bar K^{-1} difference in slopes based on the $2.5 \text{ cal mol}^{-1} K^{-1}$ entropy difference of zoisite polymorphs and ΔV_1 . The boundaries in Fig. 1 are sketched parallel to the extrapolated experimental slope of Goldsmith⁷.

The difference in Gibbs energy between clinozoisite and zoisite is readily computed from the reaction with zoisite using molar volumes of the solids^{8,9} and of H_2O (ref. 10) at 550 °C and 6.3 kbar, which is the midpoint of the uncertainty in the location of reaction (1) (Fig. 1). The volume change is $1.019 \pm 0.003 \text{ cal bar}^{-1} \text{ mol}^{-1}$ of zoisite, and the change in Gibbs energy between 6.3 and 7.3 kbar, where reaction (1) with clinozoisite is stable, is, to a good approximation, $\Delta(\Delta G_1) = \Delta V_1 \Delta P = 1,020 \pm 400 \text{ cal mol}^{-1}$ of zoisite, where the free energy uncertainty is based on the sum of the pressure uncertainties of the experimental brackets. This quantity is the Gibbs energy difference between clinozoisite and zoisite at 550 °C. The

extremely small volume difference between the phases⁹ requires that this free energy difference be virtually independent of pressure. An identical result for ΔG ($Zo \rightarrow Czo$) is obtained using the relation $\Delta(\Delta G_1) = -\Delta S_1 \Delta T$, with measured entropies of the phases^{8,10,11} and a ΔT of 45 ± 10 °C.

The most important finding of the present study is the considerably greater stability of Fe-free zoisite than Fe-free clinozoisite, despite the similarities in structure. Enami and Banno¹ showed unambiguously from field occurrences that coexisting clinozoisite and zoisite increase in Fe^{3+} with increasing temperature of metamorphism; therefore, zoisite is the relatively high entropy phase. Clinozoisite may become stable relative to zoisite, but only at a temperature well below 550 °C, in contrast to the indirect experimental deduction of equilibrium relations above 600 °C (ref. 4).

Some indication of the actual P - T stability region of Fe-free clinozoisite may be obtained from natural parageneses. The temperatures associated with the natural mineral pairs are not independently known, but may be reasonably estimated within broad limits. The lowest grade paragenesis, reported by Enami and Banno¹, from the Yoshimi area, averaged $X_{Fe} = Fe^{3+}/(Fe^{3+} + Al)$ of 0.095 for clinozoisite and 0.015 for zoisite. These greenschist grade pairs may be assigned a temperature of 350 ± 50 °C. Their highest temperature pairs, of the epidote amphibolite grade in the Iratsu locality, averaged 0.14 and 0.040, for clinozoisite and zoisite, respectively, and a temperature of 450 ± 50 °C may be assigned. A pair of compositions $X_{Fe} = 0.16$ and 0.04 was reported² from the kyanite-staurolite zone (2d) adjacent to the lowest sillimanite zone (3a) in the Moine schists of north-west Scotland (specimen 561), to which a temperature of 550 ± 50 °C may be appropriate. X-ray crystallographic studies of epidote¹² show that Fe^{3+} is almost entirely ordered in one of the three distinct octahedral M-sites. Therefore, to a good first approximation for dilute Fe^{3+} , we have for the reaction:



the equilibrium conditions:

$$\frac{-\Delta G_2}{RT} = \ln K_2 = \ln \left(\frac{1 - 3X_{Fe}^{Zo}}{1 - 3X_{Fe}^{Czo}} \right) \quad (3)$$

Figure 2 shows a plot of $\ln K_2$ versus $1/T$ with the assumed temperatures and is consistent with the broad uncertainty limits of the present experimental work. A straight line through the brackets gives an intercept of $\Delta G_2 = 0$ at ~ 200 °C, at which temperature clinozoisite becomes stable relative to zoisite. While the actual temperature of the inversion is quite uncertain because of the uncertain temperature assignments of the natural parageneses, the predicted low-temperature range is supported by several runs made in the present study, where mixtures of clinozoisite and zoisite produced growth of zoisite at 15 kbar P_{H_2O} at temperatures as low as 350 °C.

We acknowledge NSF research grants EAR 78-13675 (to J.R.G.) and EAR 81-07110 (to R.C.N.). We thank J. R. Holloway for much advice about the construction and installation of the gas pressure vessel, and P. B. Moore for discussions on epidote crystal structure. J. V. Smith provided a microprobe analysis of the experimental scolécite.

Received 13 May; accepted 17 June 1983.

- Enami, M. & Banno, S. *Miner. Mag.* **43**, 1005–1013 (1980).
- Tanner, P. W. G. *J. Petrol.* **17**, 100–134 (1976).
- Ackermann, D. & Raase, P. *Contr. Miner. Petrol.* **42**, 333–341 (1973).
- Holdaway, M. J. *Contr. Miner. Petrol.* **37**, 307–340 (1972).
- Strens, R. G. J. *Miner. Mag.* **35**, 464–475 (1965).
- Seki, Y. *Am. Miner.* **44**, 720–730 (1959).
- Goldsmith, J. R. *Am. Miner.* **66**, 1183–1188 (1981).
- Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. U.S. geol. Surv.* 1452 (1978).
- Langer, K. & Lattard, D. *Am. Miner.* **65**, 779–783 (1980).
- Burnham, C. W., Holloway, J. R. & Davis, N. F. *Geol. Soc. Am. Spec. Pap.* 132 (1969).
- Perkins, D., Westrum, E. F. & Essene, E. J. *Geochim. cosmochim. Acta* **44**, 61–84 (1980).
- Dollase, W. A. *Am. Miner.* **56**, 447–464 (1971).

Qatrania, new basal anthropoid primate from the Fayum, Oligocene of Egypt

Elwyn L. Simons

Duke University Center for the Study of Primate Biology and History, 3705 Erwin Road, Durham, North Carolina 27705, USA

Richard F. Kay

Department of Anatomy, Duke University Medical Center, Durham, North Carolina 27710, USA

Recent excavations in Egypt organized by the Geological Survey of Egypt and Duke University have recovered many fossil primate specimens, most of which come from upper levels of the Jebel Qatrani Formation (early Oligocene) including *Aegyptopithecus*, *Propliopithecus*, *Parapithecus* and *Apidium*¹⁻³. Screening at Quarry E in lower levels of the formation has also revealed a new small anthropoid described as *Qatrania wingi*. *Qatrania* is a primitive member of the Parapithecidae and possibly most closely allied to *Parapithecus fraasi*. This new species is the earliest African anthropoid known, equal in age and from the same quarry as *Oligopithecus*⁴. Its small size and details of molar structure point to a mainly frugivorous diet like some extant African prosimians and South American callitrichids.

The level of Quarry E is the lowest horizon containing abundant fossils in the Jebel Qatrani Formation and is about 50 m above the upper boundary of the underlying Qasr el Sagha Formation. The latter is of late Bartonian age, 40 Myr or older⁵. The Jebel Qatrani Formation is capped unconformably by basalt layers dated variously at 25–27 Myr⁶, so the age of Quarry E fossils is between 28 and 30 Myr (terminal early Oligocene) and 40 Myr (late Eocene). An age closer to 40 Myr is suggested by the substantial faunal turnover between lower and upper levels of the formation⁶.

Qatrania and perhaps *Oligopithecus*, another primate from Quarry E, are the oldest known anthropoids of Africa; only the poorly known and taxonomically controversial *Amphipithecus* and *Pondaungia* of the late Eocene of Burma and *Branisella* from South America are rivals as the oldest Anthropoidea worldwide⁷⁻¹³. We provisionally place *Qatrania* among Parapithecidae; primitive African anthropoids^{14,15}. *Qatrania* has derived catarrhine attributes, suggesting a relation to Old World, not New World, anthropoids.

Order, Primates Linnaeus, 1758
Family, Parapithecidae Schlosser, 1911
Genus *Qatrania*, gen. nov.

Diagnosis: *Qatrania* resembles some or all parapithecids (*Parapithecus* or *Apidium*) and is dissimilar from other Fayum taxa (*Oligopithecus*, *Propliopithecus* and *Aegyptopithecus*) in having a shallow mandible under the molar series, in lacking a P₄ protocristid, and in having a lingually placed, lingually open, P₄ talonid without entoconid crests. The distolingual placement of M₁ metaconid also is like parapithecids. Differences from any other parapithecids include very small size, and an extremely small P₄ metaconid (in other Fayum primates the paraconid is just slightly smaller than the protoconid). Additionally, M₁ has a long mesiolingually oriented paracristid supported lingually by a paraconid; other parapithecids have shorter, more mesiodistally oriented paracristids lacking paraconids¹⁶. Kay¹⁶ illustrates a small paraconid for *Parapithecus grangeri* but this cusp has proved to be a variable occurrence in this species. M₁ trigonid is large and open lingually without a premetacristid, whereas other Fayum taxa have mesiodistally short trigonids and (except for *Oligopithecus*) premetacristids¹⁶. M₃ lacks any discrete crests; each cusp stands apart and cone-like; other parapithecids have cresting.

The markedly small size and shallow roots of M₃ is a possible derived resemblance to *Parapithecus* spp.; *Apidium* spp. have M₃s equal-sized or larger than M₂. *Qatrania* lacks a buccal cingulum and a centroconid, both of which are always present on M₃s of *Apidium* spp.

Type: *Qatrania wingi*, new species.

Qatrania wingi

Diagnosis: As for genus.

Type: CGM 40240, a right mandibular ramus with P₄–M₂ and part of the root of P₃.

Horizon and locality: Collected from Quarry E, Lower Fossil Wood zone (early Oligocene)⁴. Specimen was found by Dr Scott Wing in November 1981.

Material: Type and two other specimens, a lower M₂₋₃ and an upper molar, not yet fully prepared for study.

Description: *Qatrania* is the smallest known Fayum primate, much smaller than *Apidium moustafai*, hitherto the smallest species (Table 1).

P₄–M₃ are all two-rooted. The roots of M₃ are substantially smaller and shorter than those of the other teeth (Fig. 1). In molar size, *Qatrania* has M₁ < M₂ > M₃, with M₃ the smallest tooth. The mandible is extremely shallow under the molars (Table 1, Fig. 1) and shallows slightly anteriorly.

A striking feature of P₄–M₃ crowns is the absence of any buccal cingulum except as a slight ridge mesiobuccally and distobuccally.

P₄ has a narrow oval occlusal outline. (Here and elsewhere, refer to Figs 2, 3.) Protoconid has rounded (or convex) slopes. Paraconid is absent. A small window of dentin exposed on the distolingual slope of the protoconid is the worn down remains of a metaconid. This cusp is extremely small compared with

Table 1 Dimensions (mm) of Fayum Anthropeidea

Species (and quarry)	P ₄		M ₁		M ₂		M ₃		Jaw depth under M ₁₋₂ x(s.d.)(N)
	MD x(s.d.)(N)	BL x(s.d.)(N)	MD x(s.d.)(N)	BL x(s.d.)(N)	MD x(s.d.)(N)	BL x(s.d.)(N)	MD x(s.d.)(N)	BL x(s.d.)(N)	
<i>Apidium</i>									
<i>phiomense</i> (I, M)	3.2(0.21)(37)	2.9(0.21)(37)	4.0(0.21)(67)	3.3(0.23)(63)	4.1(0.23)(59)	3.6(0.23)(57)	4.6(0.32)(49)	3.5(0.27)(48)	8.3(1.22)(35)
<i>Apidium</i>									
<i>moustafai</i> (G)	2.6(0.23)(19)	2.4(0.19)(19)	3.3(0.19)(25)	2.6(0.21)(25)	3.4(0.17)(26)	2.9(0.21)(26)	3.6(0.30)(20)	2.8(0.20)(20)	5.7(1.04)(19)
<i>Parapithecus</i>									
<i>fraasi</i> (unk., M)	3.1(–)(1)	2.9(–)(1)	3.9(–)(1)	3.5(–)(2)	4.1(–)(2)	3.5(–)(2)	4.1(–)(2)	3.0(–)(2)	7.9(–)(1)
<i>Parapithecus</i>									
<i>grangeri</i> (I)	3.8(0.26)(16)	4.0(0.20)(16)	4.6(0.23)(23)	4.3(0.32)(22)	4.7(0.22)(22)	4.3(0.33)(22)	4.6(0.35)(15)	3.7(0.24)(15)	8.9(0.70)(14)
<i>Oligopithecus</i>									
<i>savagai</i> * (E)	3.3	2.7	3.7	3.3	3.7	3.4	—	—	9.7
<i>Propliopithecus</i>									
<i>haeckeli</i> † (unk)	3.5	3.9	4.8	4.8	4.8	4.8	5.1	4.1	13.7
<i>Qatrania wingi</i>									
TYPE (E)	1.8	1.6	2.3	1.8	2.4	2.0	1.9	1.6	3.4

* From ref. 1.

† Average of left and right sides, from ref. 1.

MD, Mesiodistal length; BL, buccolingual breadth; x, mean; s.d., standard deviation; N, sample size.

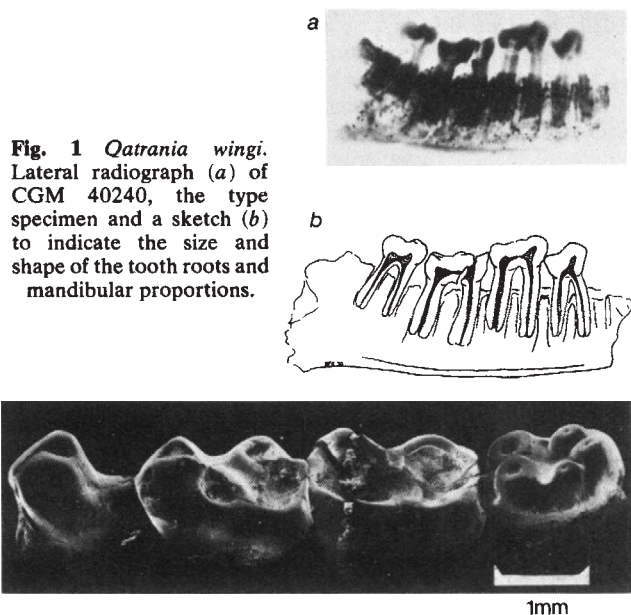


Fig. 1 *Qatrania wingi*. Lateral radiograph (a) of CGM 40240, the type specimen and a sketch (b) to indicate the size and shape of the tooth roots and mandibular proportions.

Fig. 2 *Qatrania wingi*. Composite scanning electron micrograph of P₄–M₃ (CGM 40240) occlusal surfaces from a lingual perspective.

those of any other anthropoid living or extinct and a feature resembling primitive euprimates. A round area of dentin exposed at the midline distally suggests a hypoconid. Trigonid is open lingually; a small lingually situated talonid basin is partially defined lingually by a slight distolingual cingulum.

Although heavily worn, several features of M₁ crown anatomy can be observed. The talonid is the widest part of the tooth. Trigonid is raised slightly above the level of talonid. A round exposure of dentin near the mesial end of the mesiolingually-oriented paracristid indicates a paraconid. Metaconid is distolingual to protoconid. This, and the absence of a premetacristid, leaves a large, lingually open trigonid. A strong cristid obliqua reaches mesiolingually leaving a wide hypoflexid. The distal portion of the talonid is heavily worn exposing a three-lobed, or trefoil-shaped dentin exposure, suggesting three cusps—an entoconid, hypoconid, and hypoconulid. The presence of a hypoconulid in the unworn state is also inferred from the structure of the hypocristid. This crest leads first distolingually from hypoconid, but then is notched and turns sharply distally. Such a configuration occurs only among taxa with strong hypoconulids.

M₂ is heavily worn and no aspects of its occlusal anatomy are discernable.

M₃ is well preserved. Cresting is strikingly absent. Each cusp is smoothly rounded or mamilliform. Trigonid is short mesiodistally and is much broader buccolingually than talonid. A small raised cusp (?paraconid) is present mesiobuccal to metaconid. A hypoconulid, although present, does not form a discrete heel.

The estimated body weight of *Qatrania*, based on first molar dimensions was around 300 g (Table 2)¹⁷. Estimates based on other dental dimensions give similar weights. *Qatrania*, was marmoset-sized¹⁸—or roughly one-quarter the weight of *Miopithecus talapoin*, the smallest living Old World monkey¹⁹, and one-third the estimated weight of *Apidium moustafai*, the smallest previously known parapithecoid and hitherto the smallest known catarrhine living or extinct¹⁷. For energetic reasons, this small size alone suggests that the diet of *Qatrania* did not consist of leaves as is the case in small living primates^{19,20}. Although heavily worn, it is apparent that molar crests of *Qatrania* were weak, resembling those of primarily frugivorous or gummivorous living primates and distinctly different from the trenchant molar cresting of more insectivorous species²⁰. Based on all these features, *Qatrania* probably primarily fed on fruit or vegetable gum, although insects could have provided a lesser, but important, source of nutrients.

Table 2 Body weight (g) estimates of Oligocene Anthroipoidea based on lower first molar size

	N	LN (M ₁ , MD × BL)	Body weight (±95%)
Parapithecidae			
<i>Apidium phiomense</i>	67	2.58	1,626 (1,556–1,700)
<i>Apidium moustafai</i>	25	2.15	856 (808–907)
<i>Parapithecus fraasi</i>	2	2.61	1,711 (1,638–1,787)
<i>Parapithecus grangeri</i>	23	2.99	2,975 (2,838–3,118)
<i>Qatrania wingi</i>	1	1.42	289 (265–315)
Propithecinae			
<i>Propithecus chirobates</i>	7	3.18	3,977 (3,716–4,255)
<i>Propithecus haeckeli</i>	1	3.14	3,730 (3,603–3,861)
<i>Aegyptopithecus zeuxis</i>	11	3.44	5,858 (5,652–6,072)
<i>Oligopithecus savagei</i>	1	2.50	1,448 (1,382–1,517)

Calculations are based on formulae presented by Gingerich *et al.* in Table 3 of ref. 17.

LN, Natural logarithm; ±95%, the 95% confidence interval for the estimate of the mean body weight for each species.

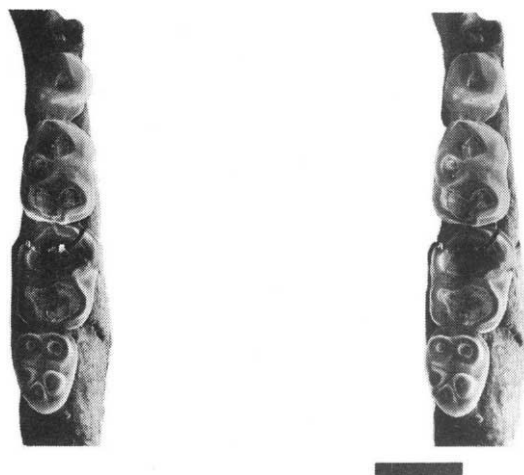


Fig. 3 *Qatrania wingi*. Stereopair scanning electron micrograph from an occlusal perspective. Scale bar, ~2 mm.

Some have suggested that parapithecids might be specially related to the New World monkeys²³, based on numerous dental and osteological resemblances^{1,2,21–23}. However, these appear to be primitive (symplesiomorphous) characters of Anthroipoidea as a whole, or obvious evolutionary parallelisms¹⁰. On the other hand, a definite special relatedness of parapithecids to Old World anthropoids is indicated by the presence in *Apidium* of such apparent synapomorphies as the absence of a vascular canal leading from the subarcuate fossa to the sigmoid venous sinus of the temporal bone²¹, a wear facet X on M₂^{10,16}, and well-developed M₁–M₂ hypoconulids¹⁰.

A few details of molar anatomy, especially that of *Parapithecus grangeri*, suggest parapithecids are in or near the ancestry of living Old World monkeys^{14,16}. The molars of *P. grangeri* show the beginnings of the cross-cresting or 'bilophodont' condition which is the hallmark of modern cercopithecoids. Reduced M₁–M₂ hypoconulids also foreshadow the condition in Old World monkeys which lack M₁–M₂ hypoconulids. The two-premolar condition of cercopithecoids is perhaps foreshadowed by *P. grangeri* which has relatively small P₂s compared with other parapithecids. Conversely, because of many primitive anthropoid characteristics of parapithecids, if these animals are the sister group of Old World monkeys, not related closely to modern apes, then many of the specialized characteristics shared by Old World monkeys and apes must have evolved in parallel^{1,24,25}. Such a list of characters includes independent loss of P₂, development of a tubular-shaped ectotympanic ossification separately in each group, as well as loss of several features of the limb bones. Whether parapithecids prove to be a separate group of primitive catarrhines or one

phylogenetically linked to Cercopithecoidea can be resolved only when more early Miocene cercopithecoids are known.

We are grateful to the Geological Survey of Egypt for the assistance they provided in all aspects of our fieldwork. This work was supported by grants from the NSF and Smithsonian Institution.

Received 4 February; accepted 15 June 1983.

1. Kay, R. F., Fleagle, J. G. & Simons, E. L. *Am. J. phys. Anthrop.* **55**, 293–322 (1981).
2. Fleagle, J. G. & Simons, E. L. *Folia primatol.* **31**, 176–186 (1979).
3. Schlosser, M. *Beitr. Paläont. Geol. Öst. Ung.* **24**, 51, (1911).
4. Simons, E. L. *Postilla, Peabody Museum, Yale Univ.* **64**, 1–12 (1962).
5. Berggren, W. A., McKenna, M. C., Hardenbol, J. & Obradovich, J. D. *J. Geol.* **86**, 67–81 (1978).
6. Bowen, B. E. & Vondra, C. F. *Ann. Geol. Surv. Egypt* **4**, 115–138 (1974).
7. Ba Maw, Ciochon, R. L. & Savage, D. E. *Nature*, **282**, 65–67 (1979).
8. Hoffstetter, R. C. *r. hebdo. Séanc. Acad. Sci., Paris, Ser. D* **269**, 434–437 (1969).
9. Gingerich, P. D. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R. & Chiarelli, A. B.) 123–138 (Plenum, New York, 1979).
10. Kay, R. F. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R. L. & Chiarelli, A. B.) 159–188 (Plenum, New York, 1979).
11. Simons, E. L. *Nature* **232**, 489–491 (1971).
12. Szalay, F. S. *Nature* **236**, 179–180 (1972).
13. Rosenberger, A. L. *Int. J. Primat.* **2**, 1–8 (1981).
14. Simons, E. L. *Postilla* **166**, 1–12 (1974).
15. Simons, E. L. & Delson, E. in *Evolution of African Mammals* (eds Maglio, V. J. & Cooke, H. B. S.) 100–119 (Harvard University, 1978).
16. Kay, R. F. *Am. J. phys. Anthrop.* **46**, 327–352 (1979).
17. Gingerich, P. D., Smith, B. H. & Rosenberg, K. *Am. J. phys. Anthrop.* **58**, 81–100 (1982).
18. Hershkovitz, P. *Living New World Monkeys (Platyrrhini)* Vol. 1 (University of Chicago, 1979).
19. Kay, R. F., thesis, Yale Univ. (1973).
20. Kay, R. F. & Covert, H. H. in *Food Acquisition and Processing in Non-human Primates* (eds Chivers, D., Wood, B. A. & Bilsborough, A.) (Plenum, New York, in the press).
21. Cartmill, M., MacPhee, R. D. E. & Simons, E. L. *Am. J. Phys. Anthrop.* **56**, 3–21 (1981).
22. Conroy, G. C. *Contrib. Primatol.* **8**, 1–134 (1976).
23. Hoffstetter, R. in *Evolutionary Biology of the New World Monkeys and Continental Drift* (eds Ciochon, R. L. & Chiarelli, A. B.) 103–122 (Plenum, New York, 1979).
24. Delson, E. & Andrews, P. in *Phylogeny of the Primates* (eds W. P. Luckett & F. S. Szalay) 405–446 (Plenum, New York, 1975).
25. Szalay, F. S. & Delson, E. *Evolutionary History of Primates* (Academic, New York, 1979).

Familial incidence of twinning

P. Parisi*†, M. Gatti†, G. Prinzi† & G. Capernat†

* School of Medicine, First University of Rome, and † The Gregor Mendel Institute of Medical Genetics and Twin Research, 00161 Rome, Italy

The frequency of twinning among women who have already borne twins, the 'repeat frequency', is significantly higher than in the general population^{1–6}. Individual propensity is not necessarily genetic in origin, but pedigree studies (for reviews see refs 1, 2) confirm that twinning is a family trait. Studies based on archives^{3,6,7} restrict this conclusion to dizygotic (DZ) twinning and the maternal side, while studies based on interviews of relatives of twins^{4,8–11} find monozygotic (MZ) twinning and the paternal side also to be involved. However, interview studies can overestimate, while archive studies can underestimate, the real frequency of twinning. We have now analysed the incidence of twinning in the families of 950 zygosity-determined, unselected twin pairs under complete ascertainment. Our results indicate that a propensity to MZ twinning, as well as one to DZ twinning, can be inherited through the maternal line, and that the two mechanisms of twinning might be related. We have also found a paternal role in DZ, but not in MZ twinning.

Our sample was drawn from the Mendel Institute's twin register, which lists ~16,000 pairs, that is, about two-thirds of all Roman twins. (Registration is free and entitles any twin pair to free medical care in some 15 specialties 5 days a week.) Selection for family history of twinning was carefully avoided: all pairs present in the institute on given days known to be days of highest attendance (mostly for the dentist or the oculist) were considered. Zygosity was in all cases determined by direct observation, detailed examination of various anthropological markers, and questionnaire procedures, with a reliability estimated at 98% through control blood-typing on ~15% of pairs. The total sample consisted of 323 MZ and 627 DZ pairs.

The pedigree information was obtained from the parents through direct interviews and then checked with other family members and on independent sources of information available at the institute. Discrepancies were clarified through supplementary interviews. Complete ascertainment was limited to two generations, so as to include the sibships of the twins, of the twins' parents and of the twins' first-degree cousins. Ascertainment bias is unlikely at these levels. The observation² that fathers, being less interested in family affairs, would be more likely to under-report single births on their side of the family, does not seem to apply to our material, as the total number of maternities was very similar in the paternal (9,623) and the maternal (9,573) sides of the family and the number of maternities per sibship was identical: 2.99 in both sides. However, it has been suggested² that, because of the difference in age at marriage of men and women, for ascertainment to be complete one should expect the father's brothers to have the same number of children as the mother's sisters, since their average length of marriage will have been the same. On this theory, the finding that father's brothers have a smaller number of children than mother's sisters in the families of both our MZ (878 versus 955) and our DZ (1,748 versus 1,843) index twins, may be taken to reflect a slight amount of under-reporting of single births in the paternal side. Our results have consequently been adjusted. Tables 1 and 2 show the results of our subsequent analysis on both the adjusted and the crude data.

For monozygotic twinning (Table 1), observed values are very close to those expected among both the twins' siblings (line 1) and paternal relatives (lines 2–5), whether crude or adjusted values are considered. On the other hand, observed values are significantly higher than expected among maternal relatives. The differences actually fail to reach significance levels among the twins' cousins (lines 7, 8), where the numbers are lower, but are highly significant ($P < 0.02$) in the mother's sibship (line 6) and, because of their general consistency, they are also highly significant ($P < 0.001$) when the totals are considered (line 9). Interestingly, the increase is found not only in the frequency of same-sex twinning ($P < 0.02$), but also in the frequency of opposite-sex twinning ($P < 0.05$ to $P < 0.02$).

For dizygotic twinning (Table 2), observed values are higher than expected among both the twins' siblings (line 1) and maternal relatives (lines 6–9). The differences are consistent throughout and reach highly significant levels not only with respect to the frequency of opposite-sex twins ($P < 0.001$), all of which are DZ, but also with respect to the frequency of same-sex twins ($P < 0.05$ to $P < 0.001$), about half of which are actually MZ. Observed values, moreover, also tend to be higher than expected among paternal relatives, and this holds true even when adjusted values are considered. The apparently inconsistent findings obtained in the father's sibship (line 2) could well be due to a mistaken attribution of one or more opposite-sex pairs to the same-sex class. (When line 2 totals are considered, in fact, the inconsistency disappears.) Otherwise, the differences are fairly consistent, although they reach significance ($P \sim 0.05$ to $P < 0.02$) only when all paternal relatives are taken together (line 5).

We believe that our procedures and sample size and characteristics make ascertainment bias unlikely. Some caution may still be desirable, however, because our results are not everywhere as clearcut as one might wish.

The finding that the incidence of twinning is very significantly higher than expected among maternal relatives of MZ as well as of DZ index twin pairs, indicates that not only the propensity to DZ twinning, but also a propensity to MZ twinning can be inherited through the maternal line, in agreement with previous family studies^{1,12–15}. Archive studies^{3,6,7} may have failed to obtain a similar result on account of possible under-reporting of twin maternities in the books. (In cases of pairs broken by discordant stillbirth or perinatal death—many times more frequent in twins than in singletons, especially in the past—the survivor is likely to have been listed as a singleton in the archives.) The fact that the frequency of twinning failed to

Table 1 Incidence of twinning among relatives of monozygotic twin pairs ($n = 323$)

Kind of relatives	No. of total maternities	No. of SS twin maternities			No. of OS twin maternities			Total no. of twin maternities		
		E	O	χ^2_1	E	O	χ^2_1	E	O	χ^2_1
1. Twins' siblings	420	3.2	2		1.6	2		4.8	4	
2. Father's sibship	1,398	12.2	18		6.7	2		18.9	20	
	<u>1,520</u>	<u>13.2</u>			<u>7.3</u>		§	<u>20.5</u>		
3. Cousins from father's brothers	878	6.6	10		3.4	6		10.0	16	
	955	7.2			3.6			10.8		
4. Cousins from father's sisters	936	7.0	1	§	3.6	5		10.6	6	
	<u>1,018</u>	<u>7.6</u>			<u>3.9</u>			<u>11.5</u>		
	3,212	25.8	29		13.6	13		39.4	42	
	<u>3,493</u>	<u>28.0</u>			<u>14.8</u>			<u>42.8</u>		
5. Total paternal side (2-4)										
6. Mother's sibship	1,427	12.6	22	7.08†	6.9	13	5.39†	19.5	35	12.47‡
7. Cousins from mother's brothers	918	6.9	8		3.5	4		10.4	12	
8. Cousins from mother's sisters	955	7.2	11		3.6	5		10.8	16	
9. Total, maternal side (6-8)	3,300	26.6	41	7.82†	14.1	22	4.52*	40.7	63	12.40‡

SS, Same sex; OS, opposite sex; E, expected values (based on age-specific, population SS and OS twinning rates, calculated separately for each sibship and within each subsample on the basis of respective mean ages); O, observed values. Underlining, values adjusted for possible under-reporting of single births in the paternal side of the family. Adjustment is based on the assumption² that the number of total maternities from father's brothers (line 3) should equal that from mother's sisters (line 8). Lines 2 and 4 were adjusted by adding to the crude values the percentage of under-reporting estimated in consequence (MZ: $(955-878)/878 = 8.77\%$; DZ: $(1,843-1,748)/1,748 = 5.43\%$). This may be an overestimate of the possible under-reporting, as it is unrealistic to assume under-reporting even in the father's sibship, so that actual values are likely to lie between crude and adjusted ones. 'Twins' siblings' exclude the index twin maternity; 'Father's sibship' includes the father; and 'Mother's sibship' includes the mother. χ^2 Analysis: values of χ^2 (for 1 d.f.) are nonsignificant unless indicated. Displayed values are significant at the following probability levels: * $P < 0.05$; † $P < 0.02$; ‡ $P < 0.001$.

§ Observed values are significantly lower than expected ones—a possible artefact due to small numbers and perhaps to some wrong attribution of one or more twin pairs to either sex-combination: in fact, the difference disappears when totals are considered.

Table 2 Incidence of twinning among relatives of dizygotic twin pairs ($n = 627$)

Kind of relatives	No. of total maternities	No. of SS twin maternities			No. of OS twin maternities			Total no. of twin maternities		
		E	O	χ^2_1	E	O	χ^2_1	E	O	χ^2_1
1. Twins' siblings	805	5.9	14	11.30‡	2.9	9	12.80‡	8.8	23	23.17‡
2. Father's sibship	2,761	24.1	35	5.00†	13.3	10		37.4	45	
	<u>2,911</u>	<u>25.4</u>		<u>3.68(*)</u>	<u>14.0</u>			<u>39.4</u>		
3. Cousins from father's brothers	1,748	12.8	17		6.3	7		19.1	24	
	1,843	13.5			6.7			20.2		
4. Cousins from father's sisters	1,902	13.9	20		6.9	9		20.8	29	
	<u>2,005</u>	<u>14.7</u>			<u>7.2</u>			<u>21.9</u>		
5. Total, paternal side (2-4)										
	6,411	50.8	72	8.95†	26.4	26		77.2	98	5.68†
	6,759	53.5		6.42†	27.9			81.4		3.44(*)
6. Mother's sibship	2,799	24.3	35	4.73*	13.3	26	12.13‡	37.6	61	14.69‡
7. Cousins from mother's brothers	1,631	11.9	19	4.24*	5.9	9		17.8	28	5.89†
8. Cousins from mother's sisters	1,843	13.5	29	18.04‡	6.7	15	10.52‡	20.2	44	28.65‡
9. Total, maternal side (6-8)	6,273	49.7	83	22.42‡	25.9	50	22.63‡	75.6	133	44.11‡

See Table 1 legend.

(*) $P \sim 0.05$.

increase among the siblings of index MZ twins, whereas it did among the siblings of index DZ twins, can be attributed to (1) the very small numbers involved in the case of MZ twins (because of smaller sample size and lower population frequency); (2) the effect of nongenetic maternal factors that influence DZ, but not MZ, twinning; and (3) apparently stronger genetic influences in the case of DZ twinning.

The finding of no increased frequency of twinning among paternal relatives of MZ index twins contradicts the results of some of the early family studies⁸⁻¹¹ and confirms the view^{2,6} that these may have been biased through under-reporting of non-twin maternities among paternal relatives. However, observations of selected families with recurrent MZ twinning through males, as have until recently¹³⁻¹⁵ been reported, cannot be merely attributed to ascertainment bias nor otherwise dismissed. It is possible that these are due to single-gene effects resulting in early developmental anomalies, and thus directly or indirectly associated to MZ twinning. If so, they should be relatively rare.

The increased frequency of twinning among paternal relatives of DZ index twins points to the existence of a paternal role in DZ twinning. Although it only reaches significance in one case and then when totals are considered, the increase is fairly consistent throughout and is not seriously affected by the adjustment for possible under-reporting. This paternal effect, presumably missed in the archive studies^{3,6,7} and overestimated in the family studies^{4,8-11} for opposite types of ascertainment bias as have been previously discussed, appears to be rather weak. That sperms can have a role in DZ twinning should not be very surprising considering that multiple ovulations in the woman are much more frequent than DZ twin maternities⁶.

The significantly increased frequency of opposite-sex (hence DZ) twin maternities among maternal relatives of MZ twins points to the existence of a relationship between the two types of twinning. This is supported by the finding that the increase of same-sex twin maternities is no lower than the increase of opposite-sex ones among the siblings, the maternal relatives and the paternal relatives of DZ index twin pairs. If the two

types of twinning are viewed as not entirely unrelated^{7-10,16}, then reports of families showing familial aggregation of twinning of both types^{1,14} become more readily understood.

Received 12 January; accepted 23 June 1983.

- Gedda, L. *Studio dei Gemelli* (Orizzonte Medico, Roma, 1951).
- Bulmer, M.G. *The Biology of Twinning in Man* (Clarendon, Oxford, 1970).
- Weinberg, W. *Pflügers Arch. ges. Physiol.* **88**, 346-430 (1901); *Arch. Rass. ges. Biol.* **6**, 322-39, 470-82, 609-30 (1909).
- Greulich, W. W. *Am. J. phys. Anthropol.* **19**, 391-431 (1934).
- Bulmer, M. G. *Ann. hum. Genet.* **23**, 31-35 (1958).
- White, C. & Wyshak, G. *New Engl. J. Med.* **271**, 1003-1005 (1964); *Am. J. publ. Hlth* **55**, 1586-1593 (1965).
- Dahlberg, G. Thesis, Univ. Uppsala (1926); *Acta genet. med. gemellol.* **1**, 80-87 (1952).
- Danforth, C. H. *J. Hered.* **7**, 195-202 (1916).
- Davenport C. B. *Z. indukt. Abstamm.-U. VererbLehre* **46**, 85-86 (1927).
- Curtius, F. & v. Verschuer, O. *Arch. Rass. ges. Biol.* **26**, 361-387 (1932).
- Slater, E. T. O. *J. Neurol. Psychiat.* **1**, 239-258 (1938).
- Lillyn, E. T. & Gindilis, V. M. *Genetika* **12**, 118-127 (1976).
- Harvey, M. A. S., Huntley, R. M. C. & Smith, D. W. *J. Pediatr.* **90**, 246-248 (1977).
- Michels, V. V. & Riccardi, V. M. *Birth Defects: Orig. Art. Ser.* **14/6A**, 201-211 (1978).
- Shapiro, L. R., Zemek, L. & Shulman, M. J. *Birth Defects: Orig. Art. Ser.* **14/6A**, 219-222 (1978).
- Boklage, C. E. *Acta genet. med. gemellol.* **30**, 167-187 (1981).

Allozymic heterozygosity and morphological variation in house sparrows

Robert C. Fleischer*, Richard F. Johnston & William J. Klitz*

Museum of Natural History, University of Kansas, Lawrence, Kansas 66045, USA

Several reports have shown that greater heterozygosity (both between individuals and between populations) is associated with lower morphological variance and asymmetry. Most previous work concerned poikilothermic organisms (for example, fish¹⁻³, butterflies⁴, lizards⁵, shellfish⁶⁻⁸, salamanders¹⁰ and plants¹¹⁻¹³). Reports concerning two homoiotherms gave conflicting results¹⁴⁻¹⁶. The report by Handford¹⁶ on a songbird, *Zonotrichia capensis*, failed to support the relationship, although these results have been questioned in the literature^{17,18}. Handford suggested that the lack of relationship in *Zonotrichia* could indicate a fundamental difference between homoiotherms and poikilotherms, reflected in their apparent differences in heterozygosity¹⁹. However, we report here on electrophoretic and morphometric studies²⁰⁻²⁵ on another songbird species (*Passer domesticus*), in which the relationship appears to be upheld. We find consistently, among four locality samples, that the class of individuals of greatest allozyme heterozygosity nearly always exhibits the lowest multivariate morphological variance, and the class of greatest homozygosity nearly always exhibits the highest

A total of 621 specimens was used in the analysis. Collections are from Manhattan, Kansas (October 1971); Point Reyes, California (December 1973)²⁵; Lawrence, Kansas (November 1978 and March 1979)²¹⁻²⁴; and Calgary, Alberta (May 1979)²³. Collection procedures involved mist-netting and occasional shooting of birds, and are outlined in greater detail elsewhere²⁰⁻²³. Tissue (kidney and liver) and serum were removed from each specimen and assayed for enzyme polymorphism using horizontal starch gel electrophoresis. Two polymorphic loci were resolved for the Point Reyes and Manhattan samples²⁵: serum transferrin (Trf-1) and kidney esterase-1 (Est-1). For Lawrence and Calgary four polymorphisms were resolved^{23,24}: Est-1, liver isocitrate dehydrogenase-2 (IDH-2), liver IDH-3, and a liver fluorescent esterase (FI-Est-1). Laboratory protocol, scoring procedures and electromorph frequencies are detailed elsewhere^{20,23}.

Individuals were grouped by locality, sex and the number of presumptive loci for which they were heterozygous. All locality samples were divided into three heterozygosity classes: 0, 1

Table 1 Variances for scores on PC1 analysis for groups of 0, 1 or 2 or more heterozygous allozyme loci

Locality	Sex	No. of heterozygous loci			P
		0	1	2	
Lawrence	Female	0.89	1.15	0.32	NS
	Male	1.00	0.57	0.46	<0.05
Calgary	Female	1.02	1.04	0.72	NS
	Male	0.62	1.08	1.12	NS
Manhattan	Female	1.23	0.86	0.55	<0.05
	Male	1.35	0.76	0.74	<0.05
Pt Reyes	Female	1.31	0.90	0.46	<0.10
	Male	1.37	0.98	0.69	<0.10

Differences among variances are tested by the Levene test. Sample sizes are the same as listed in Fig. 1. Significance for Levene test given in final column. NS, not significant.

and 2 or more loci heterozygous (sample sizes of 3- and 4-heterozygote individuals at Lawrence and Calgary were small, and were lumped with the 2-heterozygote class). We also compared each genotype and heterozygotes and homozygotes for each locus.

Specimens were measured for 14 skeletal variables^{21,26} and analysed with both univariate and multivariate techniques. Univariate and single locus results for differences in both means and variances are summarized elsewhere²³, as multivariate analysis of the variances is sufficient for illustrating the relationship. To test the null hypothesis that variances are equal among the three classes of heterozygosity, we used a multivariate extension of the Levene test^{27,28}. This test was also used by Handford¹⁶, who summarized the reasons for its use over other statistics. The Levene test compares the total sample variances (s_p^2) of two groups. Basically, it creates a new variable $y_i = (\sum_{j=1}^n (x_{ij} - \bar{x}_j)^2)^{1/2}$, where j is the number of variables, i is the number of cases, and $\bar{x}_j$ is the mean for variable j . The y_i values are averaged and compared using univariate analysis of variance (ANOVA) or t -tests.

Results of this analysis show a negative relationship between allozymic heterozygosity and morphological character variance (Fig. 1). In Fig. 1, the total variance is plotted for each class of heterozygosity. The overall trend in six of eight cases is for decreasing total variance with increasing heterozygosity. In one exception, Lawrence females, the total variance is highest in the single heterozygote class. Only for Calgary males is the total variance higher in the double heterozygote class than the zero heterozygote class.

The ratio between the total variance for the zero heterozygote class (0) and for the double heterozygote class (2) is also presented in Fig. 1. Significance levels are for the ANOVA comparing $\bar{y}_i$ of the three groups. Note that the ratio is greater than 1.0 for seven of eight comparisons, and significantly so ($P < 0.05$) for two of these (two other cases have $P < 0.10$: Lawrence females and Manhattan males). Calgary males are responsible for the case with the ratio less than 1.0, but the difference between the total variances is not significant.

Variances of scores on principal component 1 (PC1) for each class of heterozygosity were also compared using a univariate three-sample Levene test (Table 1). In previous studies of *Passer domesticus* morphometrics, as in samples used in this study, PC1 represents overall body size and explains 40-55% of the variance^{21,26,29}. The variances calculated from the PC1 scores show a pattern similar to that of total variance, and significant differences were found among three of the eight comparisons. Two of these differ from those found using mean y_i (Manhattan and Lawrence males); hence, four of the eight comparisons in some way exhibit a significant negative relationship between allozyme heterozygosity and morphological variability.

We also examined the relationship between heterozygosity and size (as assessed by the mean PC1 score for each group). For females in all four localities, size was positively related to heterozygosity. The relationship for males, however, was

* Present addresses: Department of Biological Sciences, University of California, Santa Barbara, California 93106, USA (R.C.F.); Department of Genetics, University of California, Berkeley, California 94720, USA (W.J.K.).

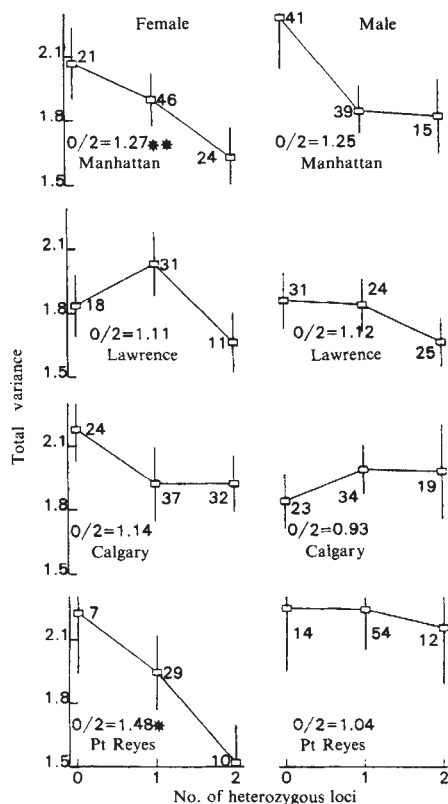


Fig. 1 Plots of the total variance ($\bar{y}_i$) for each heterozygosity class (0, 1 or 2 loci heterozygous) for each sex and locality. Vertical bars represent one standard error. Ratio O/2 is the ratio of the $\bar{y}_i$ for the 0-heterozygote class and the 2-heterozygote class. Significance levels are for an analysis of variance over the three heterozygosity groups. * $P < 0.05$; ** $P < 0.01$.

ambiguous: two samples showed an increase in size with heterozygosity, two showed a decrease. These findings will be discussed in greater detail elsewhere. Here, however, note that the usual correlation between mean and variance does not hold: groups with the largest PCI score tend to have the lowest variance.

The data show no consistent differences in body size among genotypes (that is, ff, fs, etc.). Among 24 possible comparisons, only IDH-3 for Calgary males exhibits significant heterogeneity among means on PCI ($F = 3.31$, $P = 0.04$). This proportion of significant comparisons is probably due to chance alone (that is, $1/24 < 0.05$). In addition, rank ordering of homozygote and heterozygote means (hom-hom-het or hom-het-hom) shows that heterozygote means are not predominantly interposed between homozygote means (of 21 rankable cases, 11, or 52%, show such sandwiching). Thus, differences in means among genotypic classes are probably not responsible for the trend of larger variances in homozygote classes.

We also considered the variances of single skeletal characters for each heterozygosity class. Of 112 comparisons (4 localities, male and female, and 14 skeletal variables), 14 were significantly different by Levene test, and in only 2 of these 14 were the variances of the 2-heterozygote group greater than those of the 0-heterozygote group. Furthermore, the distribution of these significant differences differed over three natural groupings of skeletal elements: limb, core and skull. Limb elements had the highest percentage of significant differences (23%) with core (9%) and skull (5%) significantly below ($G = 5.95$, $P = 0.05$).

Additionally, the 0-heterozygote group had higher variances than the 2-heterozygote group for 75% of the limb element comparisons (significantly different from 50%; $\chi^2 = 5.33$, $P < 0.025$), for 62.5% of the core element comparisons ($\chi^2 = 1.01$, $P > 0.1$), and for 52.5% of the skull element comparisons ($\chi^2 = 0.05$, $P > 0.5$). Thus, the limbs show the pattern to a

greater degree than either core or skull components. Because extremities tend to suffer environmental exposure to a greater degree than the body core (because of both surface area/volume constraints and external placement) one might expect them to be more subject to developmental instability. Thus, these results suggest that the observed relationship between allozyme heterozygosity and morphological variability may be at least partly caused by greater developmental stability or homeostasis (as in Lerner³⁴) of heterozygotes (as opposed to only additive gene effects³⁵).

Our data agree with most previous work in finding strong trends. The question remains, however, of what may ultimately cause the developmental stability of heterozygotes. Are the enzymes themselves responsible, or loci to which they are linked? Recent evidence supports the idea that intermediary metabolic enzymes may be more efficient when heterozygous⁸. This could result in greater developmental stability, and therefore, reduced morphological variance in heterozygotes. It seems unlikely, however, that the two or four polymorphic loci used in this study are entirely responsible for the trend. No single locus contributes strongly to the pattern²³. On a locus-by-locus examination, most of the comparisons (71%) show that heterozygote groups have lower multivariate variances than homozygote groups, but for only 4.2% of the comparisons are these differences significant. Significant differences are more prevalent in the multilocus examination. Therefore, it seems most likely that we are somehow representatively sampling (with our two or four loci) either the entire genome, or some segment of it that is important for developmental homeostasis (that is, metabolic enzymes). It would, however, be very difficult to accomplish such sampling on theoretical grounds^{30,31} without tight linkage disequilibrium between our marker loci and other developmentally important allozyme loci on sparrow chromosomes. Because the $2n$ chromosome number of the house sparrow is 76 (ref. 32), this would seem unlikely unless such loci are present only on a few chromosomes (or are in some other way associated). Evidence from a different study on house sparrow population genetics indicated gametic phase disequilibrium between two loci (IDH-2 and Est-1)³³. Greater understanding of the avian genome is necessary to examine this problem further. In addition, studies such as that of Koehn and Shumway⁸ are essential to uncover the actual mechanisms responsible for the phenomenon of developmental homeostasis.

We thank P. E. Lowther and W. B. McGillivray for field assistance and M. Gaines and M. R. Lein for use of laboratory facilities. M. Gaines, J. Hamrick, P. Hedrick, G. Schlager, G. Barrowclough and R. Zink provided critical commentary and J. Sauer assisted with statistics. Support of the NSF (DBS 7912412) and of The General Research Fund of the University of Kansas is acknowledged.

Received 9 December 1982; accepted 31 May 1983.

1. Mitton, J. B. *Nature* **273**, 661–662 (1978).
2. Vrijenhoek, R. C. & Lerman, S. *Evolution* **36**, 768–776 (1982).
3. Leary, R. F., Allendorf, R. W. & Knudsen, K. L. *Nature* **301**, 71–72 (1983).
4. Eanes, W. F. *Nature* **276**, 263–264 (1978).
5. Soule, M. *Evolution* **33**, 396–401 (1979).
6. Singh, S. M. & Zouros, E. *Evolution* **32**, 342–345 (1978).
7. Zouros, E., Singh, S. M. & Miles, H. E. *Evolution* **34**, 856–867 (1980).
8. Koehn, R. K. & Shumway, S. E. *Mar. Biol. Lett.* **3**, 35–42 (1982).
9. Kat, P. W. *Am. Nat.* **119**, 824–832 (1982).
10. Pierce, B. A. & Mitton, J. B. *Heredity* **73**, 250–253 (1982).
11. Knowles, P. & Mitton, J. B. *Silvae Genetica* **29**, 114–118 (1980).
12. Mitton, J. B., Knowles, P., Sturgeon, K. B., Linhart, Y. B. & Davis, M. in *Isozymes of North American Forest Trees and Forest Insects* (USDA, General Technical Report PSW-48, 1979).
13. El-Kassaby, K. *Genetics* **101**, 103–115 (1982).
14. Cothran, E. G., Chesser, R. K., Smith, M. H. & Johns, P. E. *Evolution* **37**, 282–291 (1983).
15. Smith, M. H., Chesser, R. K., Cothran, E. G. & Johns, P. E. *Antler Development in Cervidae* (ed. Brown, R. D.) (Caesar Kleberg Wildlife Research Institute, Texas A & I University, 1982).
16. Handford, P. *Nature* **286**, 261–262 (1980).
17. Eanes, W. F. *Nature* **290**, 609–610 (1981).
18. Handford, P. *Nature* **290**, 610 (1981).
19. Nevo, E. *Theor. population Biol.* **13**, 121–177 (1978).
20. Klitz, W. thesis, Kansas Univ. (1972).
21. Johnston, R. F. & Fleischer, R. C. *Auk* **98**, 503–511 (1981).
22. Fleischer, R. C. & Johnston, R. F. *Nature* **298**, 747–749 (1982).
23. Fleischer, R. C. thesis, Kansas Univ. (1983).
24. Fleischer, R. C. *Evolution* (in the press).

25. Johnston, R. F. & Klitz, W. J. in *Granivorous Birds in Ecosystems* Vol. 12 (eds Pinowski, J. & Kendeigh, S. C.) ch. 2 (Cambridge University Press, 1977).
26. Johnston, R. F. & Selander, R. K. *Evolution* **21**, 1-28 (1971).
27. Van Valen, L. *J. theor Biol.* **45**, 235-247 (1974).
28. Van Valen, L. *Evolutionary Theory* **4**, 33-43 (1978).
29. Johnston, R. F. *Syst Zool.* **22**, 219-226 (1973).
30. Mitton, J. B. & Pierce, B. A. *Genetics* **95**, 1043-1054 (1980).
31. Chakraborty, R. *Genetics* **98**, 461-466 (1981).
32. Bulatova, N. S., Radjabli, S. I. & Panov, E. N. *Experientia* **28**, 1369-1371 (1972).
33. Cole, S. R. & Parkin, D. T. *Biol. J. Linn. Soc.* **15**, 13-22 (1981).
34. Lerner, I. M. *Genetic Homeostasis* (Oliver & Boyd, Edinburgh, 1954).
35. Chakraborty, R. & Ryman, N. *Genetics* **103**, 149-152 (1983).

Cell cycle control by timer and sizer in *Chlamydomonas*

Lorraine Donnan & Peter C. L. John

Department of Botany, The Queen's University of Belfast, Belfast BT7 1NN, UK

Conservation of cell cycle control mechanisms is indicated by the presence of functionally homologous division control genes in unrelated yeasts¹ and by the nonspecific action of oncogenes^{2,3}, but it remains uncertain what property of a growing cell results in the initiation of events leading to division. Response to a critical size is indicated by the longer growth period of smaller cells prior to division⁴⁻⁶, which is consistent with deferment of division events until a minimum size is attained; however, in the same cell types faster growing cells are larger⁷⁻⁹ and this is more easily explained if division follows a timed period during which faster growing cells grow more, as is postulated for mammalian cells^{10,11}. Therefore, either time- or size-dependent controls might be the sole significant mechanism^{12,13}; we report here, however, that both controls do function in *Chlamydomonas* since cycle duration is under timer control and cell size determines the number of division rounds committed at the end of each cycle, and hence whether 2, 4, 8 or 16 daughter cells are formed.

Photosynthesizing cells of *Chlorella*^{14,15} and *Chlamydomonas*^{16,17} attain a commitment to divide before starting DNA synthesis and thereafter they will complete division without requirement for growth, but it has not been established whether the time of commitment is controlled by attainment of a critical size or by a timer.

Evidence for a timer is provided by the commitment times of cells growing at unchanged temperatures of 20 °C or 30 °C, and of cells transferred between these two temperatures. There was a 2.5-fold difference in growth rate between cells remaining at 30 °C and cells transferred to 20 °C (indicated by protein doublings of 4.4 h and 11 h respectively), and other cultures were intermediate between these growth rates (Fig. 1a), but in all four cultures there was a similar rate of progress to a state of commitment to divide, which was tested by incubating samples in darkness. Darkness removes the external energy supply and prevents further growth and new commitments, but cells that are already committed will within 2 h enter S phase and continue through division. The mid-time of first commitment to produce two daughter cells occurred between 8.5 h and 10 h after the start of growth in daughter cells at each temperature (Fig. 1b,c). As the temperature-compensated measurement of a time interval is the prime characteristic of a biological timer¹⁸, we conclude that commitment to divide is under the control of a timer.

Slower growth caused by energy limitation does, however, extend the timed period, and study of synchronously dividing cultures growing at 25 °C under a range of light intensities has shown that over a fivefold range of protein doubling times between 5 h and 25 h, the average time at which cells become committed to division is related to their growth rate (Fig. 2a) according to the equation:

$$\text{Time of commitment (h)} = 5 + 0.75 \times \text{doubling time (h)} \quad (1)$$

Operation of a timer, rather than a size control of commitment,

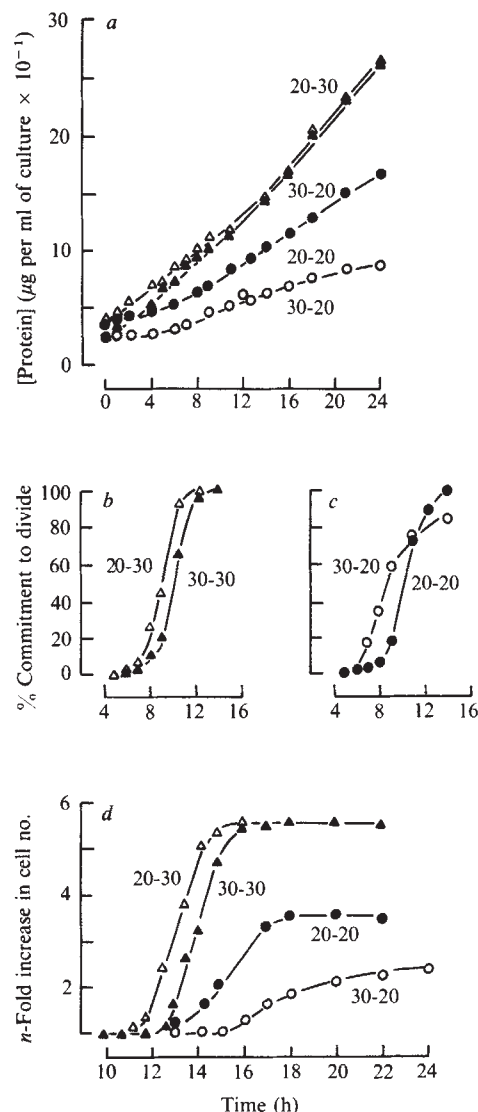


Fig. 1 Effect of temperature on: a, growth; b, c, time of commitment to first division; and d, formation of daughter cells, in *Chlamydomonas reinhardtii* 137C mt⁺. Cells were grown photoautotrophically in mineral salts medium and were synchronized in batch culture by repeated cycles of 14 h illumination and 10 h darkness in high salt medium as described previously²⁷. Cells were grown for four cell cycles at constant temperature and then at time 0, cells previously grown at 20 °C were either: ●, maintained at 20 °C; or △, transferred at time 0 from 20 °C to 30 °C by immersion in a water bath; and cells previously grown at 30 °C were either: ▲, maintained at 30 °C; or ○, transferred from 30 °C to 20 °C by immersion in a transparent water bath regulated by a Churchill temperature controller. All cultures were adjusted to the same A₆₈₀ at time 0 and were aerated with 0.5% v/v CO₂ in air and illuminated in parallel, at 200 μeinstein m⁻² s⁻¹ of light in the range 400–700 nm from warm white fluorescent lights on both sides of the culture. Cells adapted to 30 °C had a higher content of photochemical apparatus and at time 0 were therefore at a lower protein concentration per ml of culture. Commitment to divide was determined in samples of culture transferred to darkness and aerated at the same temperature, by microscopic inspection of at least 500 potential mother cells per sample for the presence of daughter cells. Progress of cytokinesis shown in d, was determined in cells that were not subjected to darkness, by counting completely formed daughter cells prior to their release from the mother cell envelope. Total protein was estimated in samples of cells that were immediately pelleted at 4,000g and suspended in cold 25 mM Tris-HCl, pH 8.0 and stored at -20 °C. Samples were thawed and the cells broken by passage through an Aminco French pressure cell at 1.4 × 10⁷ kg m⁻², then protein was precipitated by 85% (v/v) acetone and dissolved in 0.5 M NaOH for assay by the Lowry method²⁷.

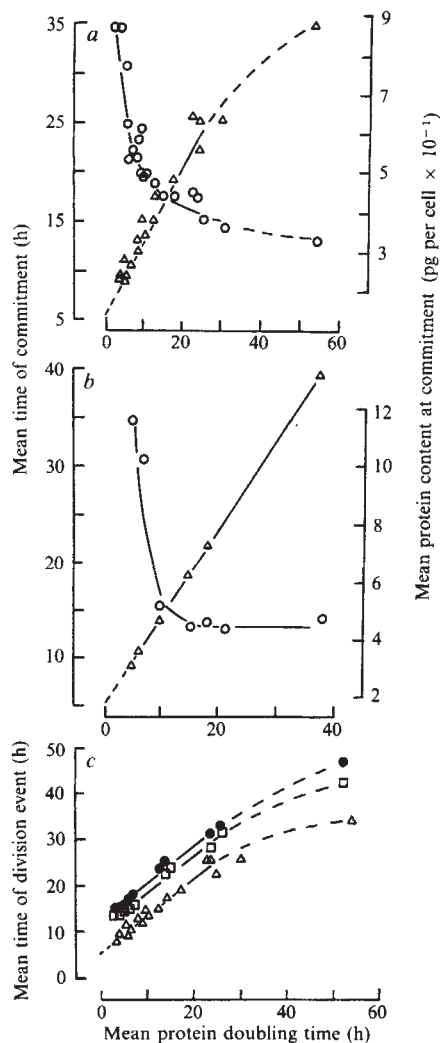


Fig. 2 Effect of growth rates limited by energy on: Δ , time of commitment to first division; $\circ$, protein content at commitment; $\square$, mid-time of final cytokinesis; and $\bullet$, time of completion of cytokinesis. *a, c*, Cells that were formed by division during a preceding 10-h dark period; *b*, cells that were formed by division of illuminated mother cells and were in their second cell cycle under continuous illumination. Times throughout are hours after the start of growth in autonomous daughter cells, that in *a* occurred at the start of illumination and in *b* occurred at the mid-time of the final cytokinesis, which was estimated as the time at which the cell number was 0.75 of the final cell number at completion of the division round. Mid-time of first commitment to division was determined by a method that accommodates slow-growing cultures, by measurement of the time at which the committed cell number became 1.5 of the initial number. This marks the time at which 50% of cells have passed their first commitment to divide, since cells first become committed to produce two daughters, although they can soon subsequently commit successively to further doublings of daughter number if size permits and growth continues (see Fig. 4 and text). This first commitment is identical with that measured in Fig. 1. Time of completion of cytokinesis was the time at which increase in autonomous cell number ceased. In *a* the solid line passing through the points indicating mean protein content at commitment has been calculated, using equation (1), from the mean protein content of 21 pg per cell at time 0. Amounts of protein are in weights of bovine serum albumin, equivalent to Lowry assay-reacting material. Mean protein content at time of first commitment was determined by measurement of protein accumulation, as described in Fig. 1 legend. Doubling times for protein were derived from at least 10 estimations of protein content during each cell cycle. Growth during the cell cycle was taken to be exponential²⁷ and doubling times indicate mean growth rate from the start of daughter cell growth to the time of each event. Different growth rates were achieved by illuminating cultures at different light intensities between 20 and 260 $\mu\text{Einstein m}^{-2} \text{s}^{-1}$ in the range 400–700 nm.

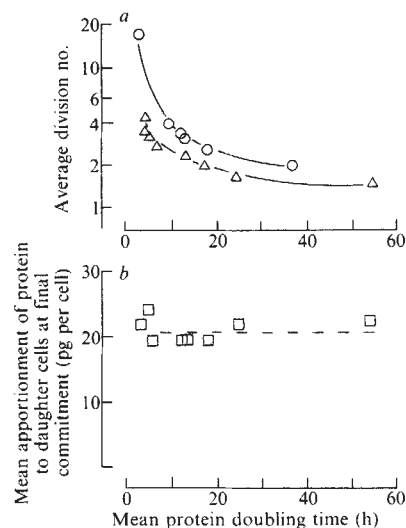


Fig. 3 Effect of energy-limited growth rates on: *a*, average division number; and *b*, average apportionment of mother cell protein to daughter cells at the time of final commitment to divide, in *Chlamydomonas*. During the experiment, cells were subjected to continuous illumination. In *a* the average division number was estimated by comparing cell number before division with that when division was complete. Division number was estimated in: Δ , cells that were formed during a dark period before beginning the cell cycle leading to the division that was investigated and: $\circ$, in cells that were formed in continuous light before the cell cycle that was investigated. In *b* the mean apportionment ($\square$) of mother cell protein at the mid-time of final commitment to divide was determined by estimating the protein per ml of culture at the time when committed cell number was 0.75 of the final cell number, and dividing this protein level by the number of daughter cells eventually formed. Protein level, cell number and commitment to divide were estimated as described in the legends to Fig. 1 and 2.

is indicated by the severalfold range of mean cell sizes at which populations with different growth rates reached commitment (Fig. 2*a,b*). The possibility that time of commitment is determined by attainment of a critical size that is set at higher levels at faster growth rates was excluded by studying the larger daughter cells that formed when mother cells were illuminated throughout division (Fig. 2*b*). Such mother cells continued growing during the final S, mitosis and cytokinesis phases and produced larger daughter cells. The size difference as measured by protein content was accentuated by the degradation of protein in cells dividing in darkness and the difference was greater in illuminated cells that were growing more quickly, but the mean protein content of daughters formed in darkness was between 50 and 70% of that in illuminated cells. Larger cells would be expected to reach a critical size sooner and so to commit earlier, but over a more than fivefold range of growth rates both populations, when growing at the same rate, reached commitment at the same time, in accordance with equation (1) (Fig. 2*a,b*). The levelling-off of curves for protein content at commitment, seen at slow growth rates, does not indicate a requirement for a critical minimum size. Because this form of curve is predicted by the timer equation (1), and commitment to division at an inadequate size at very slow growth rates can be avoided without a mechanism responsive to size because with very low energy supply, cells cease dividing. However, the possibility that at very slow growth rates, in other conditions, size does control commitment requires further study.

An implication of equation (1) is that faster growth rates (doubling times shorter than 20 h) will result in cells that have more than doubled in size between formation and commitment. A theoretical basis is therefore provided for the correlation of higher division numbers with faster growth rates that has been observed in *Chlorella*¹⁹ and *Chlamydomonas*^{20,21}.

It has been proposed that higher division numbers in algae dividing by multiple fission are a consequence of the

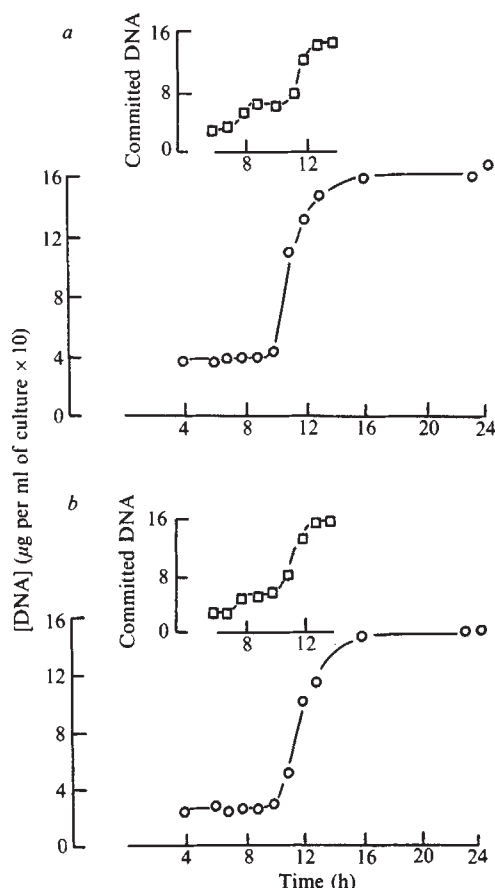


Fig. 4 Effect of average cell size: *a*, small; *b*, large; on the increase in DNA (○) and inset, commitment to DNA synthesis (□), in cells of *Chlamydomonas* growing at the same rates. Time is relative to the start of illumination, and corresponds to the duration of growth after daughter cells became autonomous. The cells had previously been grown at 25 °C in synchronizing conditions, with 14-h periods of illumination at $200 \text{ einstein m}^{-2} \text{ s}^{-1}$ in the range 400–700 nm and 10-h dark periods. When daughter cell release was complete, a culture was subjected to differential centrifugation to obtain populations of small and large average cell size, which were then adjusted to the same culture turbidity as that before fractionation and then illuminated and aerated in parallel in the same conditions as for earlier growth. Note that the 3-h separation between first and second commitments to DNA duplication allows resolution of separate commitments, and this is consistent with evidence for the occurrence of separate complete rounds of DNA replication obtained by heavy isotope labelling²⁰ and microassay of DNA in individual nuclei²⁸, however the resulting DNA duplications follow each other more promptly than do the individual commitments and are not resolved due to limitations of synchrony. DNA content was estimated by diphenylamine reaction²⁷ and commitment to DNA synthesis was determined by assay of the final DNA level in darkened subcultures.

physiological state of faster-growing cells²², but the present study shows that size of mother cells and not growth rate determines division number, because the larger daughter cells, which reached commitment at larger sizes than the smaller daughters (Fig. 2*a, b*), adopted higher division numbers at equivalent growth rates over a 10-fold range of doubling times (Fig. 3*a*). Further evidence that mother cell size is the determining factor for division number, comes from the centrifugal selection²³ of large and small cells from a single population of newly released daughter cells. When grown in identical conditions the two populations showed almost identical protein doubling times of 4.9 and 4.7 h, respectively, for the populations of large and small average size. In both cultures, commitment to doubling of DNA occurred at 9 h (Fig. 4) and mother cells subjected to darkness at 9 h produced predominantly two

daughter cells. A further commitment occurred in both cultures at 12 h and resulted in a fourfold overall increase in DNA (Fig. 4) and a predominance of four daughter cells. Only in the population of larger mean cell size did a significant number of cells undertake a further commitment, with a consequent >6.2-fold increase in DNA (Fig. 4), and most daughters formed in eights. The two populations were grown in identical environmental conditions and showed nearly identical growth rates, therefore larger size is clearly the factor that allows progression through a greater number of successive commitments. An indication of the minimum mass requirement for a further commitment to divide was obtained by comparing mean mother cell mass at the mid-time of final commitment with mean division number, in cells growing at different rates (Fig. 3*b*). Clearly, there is a mean apportionment of about 20 pg protein to committed daughter cells over the whole range of energy-limited growth rates that have studied. The simplest explanation is that further successive commitments occur while the apportionment of mother cell mass to committed daughter cells remains greater than 27 pg protein; in consequence the minimum apportionment at final commitment is 13.5, the maximum is 27 and the mean is 20 pg of protein.

The post-commitment period (from first commitment through DNA syntheses and mitoses to completion of divisions) has different properties as its events are not dependent on concurrent growth and it is less temperature-compensated. In cells fully adapted to the prevailing temperature the period between mean time of first commitment and the completion of cytokinesis was 5.5 h at 30 °C and 7 h at 20 °C (Fig. 1*a, d*), indicating a Q_{10} of 1.3 in contrast with the pre-commitment Q_{10} of 1. Sensitivity to temperature change was shown by the more than doubling in duration of the post-commitment period in cells transferred from 30 °C to 20 °C (Fig. 1*d*). However, the duration of the post-commitment period was relatively unaffected by growth rate, as faster biosynthesis allows faster progress through events but is compensated by the need to perform a greater number of divisions. Constancy of the post-commitment period, at mass doubling times between 5 h and 25 h at 25 °C, allows the duration of the cycle, from the start of growth in autonomous daughter cells to the mid-time of their subsequent final cytokinesis (Fig. 2*c*), to be described by changing the constant in equation (1):

$$\text{Mean generation time (h)} = 11 + 0.75 \times \text{doubling time (h)} \quad (2)$$

The constancy of the post-commitment period indicates that the first commitment to divide is a major rate-limiting step in the cell cycle, analogous to 'start' in other cells¹, as well as marking a biochemical breakpoint between events that are temperature-compensated and extended by slow growth (equation (1)), and events that are more temperature-sensitive but continue in the absence of growth. These biochemical differences result in the synchronization of divisions under periodic illumination, because cells in darkness do not undertake further commitments but do complete divisions that are in progress, and so on re-illumination all cells are in pre-commitment phase and synchronously initiate the commitment timer.

The detection of the two different timer periods with different properties indicates that the *Chlamydomonas* cell cycle is not under the control of a single endogenous circadian oscillator timer¹⁸ running continuously through successive cycles, as has been postulated²⁴. The commitment and post-commitment timer periods each start on completion of the other and without reference to the timing of previous cell cycles. They therefore fall into the class of hourglass timers²⁵ because they measure a single period of time following a starting stimulus, as a flow of sand is started by inversion of an hourglass. By studying an organism in which cell cycle duration and division number can be independently varied, we have demonstrated both timer- and size-dependent controls, thus providing support for general models of cell cycle control that accommodate both phenomena^{7,13,26}.

We thank Miss Eileen Carvill for assistance and the SERC for support. L.D. thanks the DENI for support.

Received 15 April; accepted 30 June 1983.

1. Beach, D., Durkacz, B. & Nurse, P. *Nature* **300**, 706-709 (1982).
2. Logan, J. & Cairns, J. *Nature* **300**, 104-105 (1982).
3. Colby, W. W., Chen, E. Y., Smith, D. H. & Levinson, A. D. *Nature* **301**, 722-725 (1983).
4. Fantes, P. A. *J. Cell Sci.* **24**, 51-67 (1977).
5. Johnston, G. C., Pringle, J. R. & Hartwell, L. H. *Expl Cell Res.* **105**, 79-98 (1977).
6. Shields, R., Brooks, R. F., Riddle, P. N., Capallero, D. F. & Delia, D. *Cell* **15**, 469-474 (1978).
7. Fantes, P. A. & Nurse, P. *Expl Cell Res.* **107**, 377-386 (1977).
8. Jagadish, M. N., Lorincz, A. & Carter, B. L. A. *FEMS microbiol. Lett.* **2**, 235-237 (1977).
9. Tovey, M. & Brouty-Boye, D. *Expl Cell Res.* **101**, 346-354 (1976).
10. Smith, J. A. & Martin, L. in *Cell Cycle Controls* (eds Padilla, G. M., Cameron, I. L. & Zimmerman, A.) 43-60 (Academic, New York, 1974).
11. Brooks, R. F., Bennett, D. C. & Smith, J. A. *Cell* **19**, 493-504 (1980).
12. Brooks, R. F. *Soc. exp. Biol. Semin. Ser.* **10**, 35-61 (1981).
13. Pardee, A. B., Shilo, B. Z. & Koch, A. L. in *Hormones and Cell Culture* (eds Sato, G. & Ross, R.) 373-392 (Cold Spring Harbor Laboratory, New York, 1979).
14. Tamiya, H., Iwamura, T., Shibata, K., Hase, E. & Nihei, T. *Biochim. biophys. Acta* **12**, 23-40 (1953).
15. Wanka, F. *Arch. Mikrobiol.* **34**, 161-188 (1959).
16. Howell, S. H., Blaschko, W. J. & Drew, C. M. *J. Cell Biol.* **67**, 126-135 (1975).
17. Spudich, J. L. & Sager, R. *J. Cell Biol.* **85**, 136-145 (1980).
18. Daan, S. in *Biological Timekeeping, Soc. exp. Biol. Semin. Ser.* (ed. Brady, J.) 11-32 (Cambridge University Press, London, 1982).
19. Molloy, G. R. & Schmidt, R. R. *Biochem. biophys. Res. Commun.* **40**, 1125-1133 (1968).
20. Kates, J. R., Chiang, K. S. & Jones, R. F. *Expl Cell Res.* **49**, 121-135 (1968).
21. Mihara, S. & Hase, E. *Pl. Cell Physiol., Tokyo* **12**, 225-236 (1971).
22. Poyton, R. O. *J. Bact.* **113**, 203-211 (1973).
23. John, P. C. L., Lambe, C. A., McGookin, R., Orr, B. & Rollins, M. J. *J. Cell Sci.* **55**, 51-67 (1982).
24. Edmunds, L. N. & Adams, K. J. *Science* **211**, 1002-1013 (1981).
25. Lees, A. D. *J. Insect Physiol.* **19**, 2279-2316 (1973).
26. Nurse, P. *Nature* **286**, 9-10 (1980).
27. Rollins, M. J., Harper, J. D. I. & John, P. C. L. *J. gen. Microbiol.* **129**, 1899-1919 (1983).
28. Coleman, A. W. *J. Phycol.* **18**, 192-195 (1982).

Cadmium and mercury nephrotoxicity

J. K. Nicholson

Toxicology Unit, Department of Pharmacology, School of Pharmacy, Brunswick Square, London WC1N 1AX, UK

Marion D. Kendall

Department of Anatomy, St. Thomas's Hospital Medical School, London SE1 7EH, UK

D. Osborn

Institute of Terrestrial Ecology, Monks Wood Experimental Station, Huntingdon PE17 2LS, UK

Despite increasing attempts to control environmental pollution, changes in the distribution and bioavailability of toxic metals like mercury and cadmium are still occurring. Apart from natural processes, other contributory factors include the gradual spread of industrialization, the use of sewage sludge as a fertilizer and the acidification of Northern Hemisphere ground-water. Animals (including man and domestic varieties) can accumulate harmful concentrations of toxic metals¹⁻⁵. We therefore looked for damage to the kidneys in seabirds contaminated with mercury and cadmium and made comparisons with kidneys from three other groups of animals: seabirds from an uncontaminated colony, metal-dosed birds and metal-dosed mice. We report here that, comparing all these groups of animals, individuals with comparatively high levels of metals had nephrotoxic lesions of a similar type and severity. Moreover, the metal concentrations at which damage began and at which biochemical changes could be detected were below those presently considered as relatively safe for humans by the World Health Organization.

Previous studies have shown that many animals accumulate high concentrations of toxic metals¹⁻⁵. Some animals are polluted with metals from anthropogenic sources, while others appear to be contaminated from natural sources. On the basis

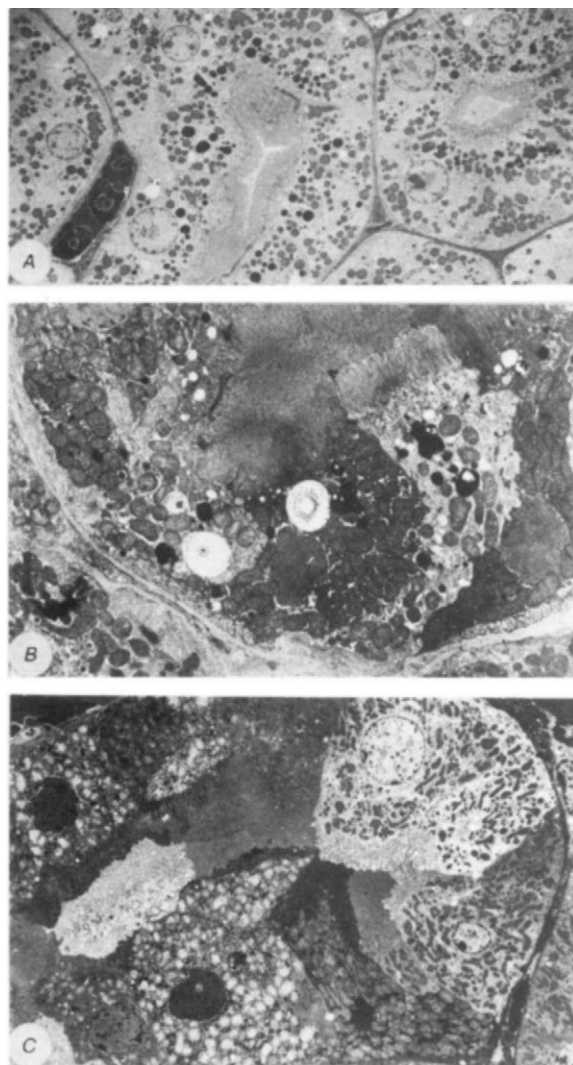


Fig. 1 Electron micrographs of: *A*, proximal tubules and kidney cortex in normal wild starling (*Sturnus vulgaris*). All the cells are of an even electron density. $\times 1,275$. *B*, Proximal tubules of Manx shearwater (*Puffinus puffinus*), Isle of St Kilda, showing the range of electron density of the cells possibly associated with high kidney metal concentrations. The most electron-dense cells have pyknotic nuclei, numerous enlarged mitochondria, and contain vacuoles and heterogeneous dense bodies. The basolateral membrane interdigitations are abnormally clearly seen (in contrast to normal birds). $\times 3,550$. *C*, Kidney cortex of mouse, 3 days after the last of 13 subcutaneous doses of CdCl_2 ($0.14 \mu\text{moles}$ per injection) given at 3-day intervals. The proximal tubule, as in the wild Manx shearwater (*B*) contains cells of varying electron density. The swollen, large mitochondria give the cells a foamy appearance (cloudy swelling) under light microscopy. Recently regenerated cells appear electron-lucent and necrotic cells have very electron-dense cytoplasm and microvilli. $\times 1,900$. **Methods:** Kidneys were fixed *in situ* after anaesthetization with Sagatal ($0.01 \text{ ml per g body weight}$), excised for further fixation (3 h) and processed routinely through 2.5% glutaraldehyde and 1% osmium tetroxide. For *B*, the 1% osmium tetroxide step was omitted.

of broad ecological judgements (for example, apparent breeding performance of colonies of seabirds containing different metal concentrations), it could be concluded that animals exposed to high levels of metals for a long time have evolved mechanisms to minimize the effects on breeding potential. This seemed to be the case with seabirds on St Kilda⁶. This does not, however, exclude the possibility that the animals are affected by the metals. Therefore, we have studied, at the ultrastructural level, the effects of cadmium and mercury on the kidney. Furthermore, as there are potential difficulties in extrapolating between species, we also studied laboratory birds and mice which had

Table 1 Comparison of kidney pathology in Cd- and Hg-dosed A₂G Swiss mice and starlings with that of two species of St Kilda seabirds

Range of kidney metal concentrations (mg per kg dry wt)	Metal-dosed animals				St Kilda seabirds (Fulmar petrel and Manx shearwater) [Cd] = 60–480 [Hg] = 3–62 n = 9
	A ₂ G mice		Starlings		
	[Cd] = 110–260 n = 12	[Hg] = 70–160 n = 4	[Cd] = 95–240 n = 7	[Hg] = 25–60 n = 5	
Renal corpuscles					
Podocyte vacuolation	++	+	++	++	++
Mesangial matrix proliferation	++	+	+	+	+
Endothelial cell swelling/proliferation	++	—	—	—	—
Proximal tubule					
Cell necrosis	+++	+++	+++	+++	++
Nuclear pyknosis	+++	++	++	++	++
Mitochondrial swelling	+++	+++	+++	+++	++
Tubulorrhexis	+	+	++	+	—
Regenerative activity*	+++	++	+	+	+
Hydropic changes in distal tubules	+	++	+	+	+
Cortical inflammation (chronic)	++	+	++	+	—
Cellular debris in distal nephron lumen	+++	+++	+++	+++	++

For the St Kilda seabirds (the Fulmar petrel, *Fulmaris glacialis*, and the Manx shearwater, *Puffinus puffinus*), $n = 9$ refers to a previous study⁷; for EM observations, $n = 3$ for each species. In the case of the A₂G mice and starlings, 10–13 subcutaneous injections of 0.14 μmol (mice) and 0.28 μmol (starlings) Cd²⁺ or Hg²⁺ were given in 0.1 and 0.2 ml of saline at 3-day intervals. Three days were allowed to elapse after final dosing before the animals were killed. — Indicates no detectable abnormality; +, slight abnormality; ++, moderate changes; +++, severe changes. Although Hg and Cd had broadly similar effects on the kidney, it is probable that Cd is the more important environmental hazard as its half life in tissues is greater than that of Hg.

* Assessed subjectively from numbers of partially undifferentiated cells present.

been dosed with these metals. We hoped this study would help to elucidate the mechanisms underlying metal damage to kidneys. At the whole animal level, such effects on the kidney would probably manifest themselves differently in different animals. For example, they could be more important in free-living than in laboratory animals, as the former are exposed to hazards such as bad weather, food shortage and disease, which would exacerbate the effects of organ damage.

Seabirds from the north-west UK colony of St Kilda (mean kidney Cd concentration 100–200 mg per kg dry tissue and mean kidney Hg 5–13 mg per kg dry tissue)^{5,7} had structural lesions of a nephrotoxic type that were visible by light microscopy (LM) in 1 μm araldite-embedded sections. Under LM^{8,9}, patchy necrosis of the proximal tubules and some glomerular damage was seen, and in the worst affected birds the majority of the proximal tubule profiles contained one or more necrotic cells whose cytoplasm was dark with pale, swollen mitochondria. Such cells were rarely seen in sections of kidney in birds from the relatively uncontaminated Isle of May colony⁸ (mean kidney Cd 5–15 mg per kg dry tissue and mean Hg 2 mg per kg dry tissue¹⁰). We examined the St Kilda seabird kidneys by electron microscopy (EM) to obtain more detailed characteristics of the lesions (Table 1).

Both the LM and EM lesions seen in seabirds (Table 1; Fig. 1) were also present in birds dosed with Cd or Hg. In the dosed birds, kidney metal levels were comparable to those found in St Kilda birds, supporting the view that the seabird kidney lesions were metal-induced and not a result of some other factor. Further similarities between the dosed starlings and the seabirds included elevated Zn concentrations⁸ and the presence of a 'metallothionein-like' cadmium and zinc binding protein in the tissues^{8,11}.

Electron microscope studies (conventional preparative routines) of Cd-dosed Swiss A₂G mice, whose tissue Cd concentrations were similar to those of the wild and experimental birds, again showed similar nephrotoxic lesions (Table 1; Fig. 1C). Cd-dosed mice also had elevated kidney, liver and pancreatic Zn concentrations^{8,12}. This disturbance of a naturally occurring element led us to examine other elemental differences by quantitative X-ray microanalysis of control mice and Cd-dosed mice having cells of normal appearance and cells in various stages of degeneration, respectively. Early results^{13,14} from this study show that considerable changes in proximal tubule cell elemental composition occur, the main features

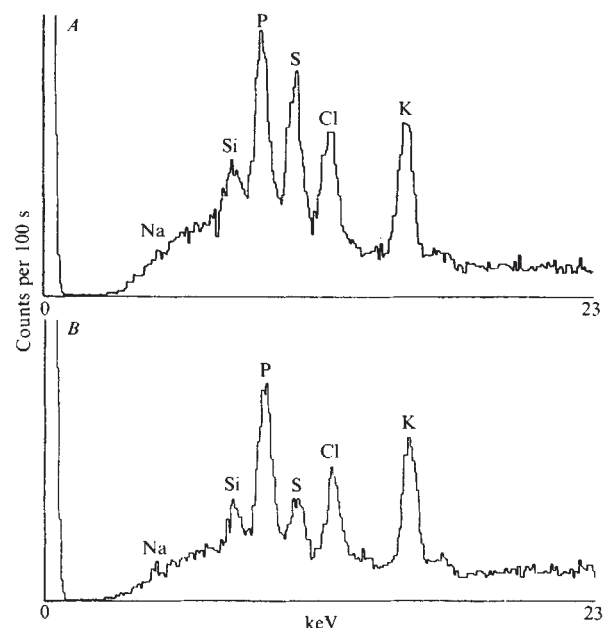


Fig. 2 Typical X-ray emission spectra (Kevex detector with a Link 290 Multichannel analyser fitted to an EMMA-4) for Cd-dosed (A) and control (B) mice. Spectra for distal tubule mitochondria are shown. For Na, Mg, P, S, Cl, K and Ca the significant concentration changes in dosed mice were as follows: Na↓ ($P < 0.001$), K↓ ($P < 0.01$), Mg↓, P↓, Cl↓ ($P < 0.1$) in the proximal tubule mitochondria, and Mg↑ ($P < 0.1$) in distal tubule mitochondria and cytoplasm. In addition, there was a highly significant increase in S↑ ($P < 0.001$) in both mitochondria (156 ± 12 mmol per kg freeze-dried mass, controls; 334 ± 32 mmol per kg freeze-dried mass, Cd-dosed) and cytoplasm (58 ± 4 mmol per kg freeze-dried mass, controls; 126 ± 12 mmol per kg freeze-dried mass, Cd-dosed) in distal tubules. Labels refer to K α emissions.

being a lowering of many elemental concentrations in the proximal tubules after Cd dosing, consistent with a loss of ionic regulation in the cells due to changes in membrane permeability, and a highly significant rise in sulphur levels in the distal tubule mitochondria and cytoplasm (Fig. 2). Recently, immunohistochemical localization of sulphur-rich metallothionein has been demonstrated in the distal and collecting tubules in rats¹⁵.

The apparent loss of ionic control in the proximal tubules is consistent with the known inhibitory effect of Cd on ($\text{Na}^+ + \text{K}^+$)ATPase¹⁶, perhaps related to the structural alterations in mitochondria seen in the nephrotoxic lesions described here.

Although the animals we studied were outwardly healthy, and while the regenerative capacity and functional reserve of the kidney is probably substantial, we suggest that our findings (and those of others¹⁷) indicate that a review of the generally accepted 'safe' levels of Cd in animal tissues is now required. This is because (1) we found renal lesions and associated biochemical changes in birds and mammals at Cd concentrations of the same order as found in western adult human populations^{18,19}; and (2) we observed potentially irreparable kidney damage (tubulorrhexis) and inflammation at kidney Cd concentrations which are below 200 mg per kg wet weight cortex (that is, >600 mg per kg dry weight cortex), a level normally considered by the WHO²⁰ and others²¹ to have no serious effects on kidney function in man. More research is needed on the significance of Hg-induced lesions because of the different proportions of organic forms of Hg found in various species.

We thank NERC and RTZ Services Ltd for financial support; Dr T. C. Appleton and Dr A. Warley for their advice and the use of the X-ray microanalysis facilities in the Department of Physiology, University of Cambridge, UK; and St Thomas's Hospital Research (Endowments) Fund for some of the costs.

Received 28 January; accepted 14 June 1983.

1. Friberg, L., Nordberg, G. F. & Vouk, V. B. *Handbook on the Toxicology of Metals* (Elsevier, Amsterdam, 1979).
2. Waldbott, G. L. *Health Effects of Environmental Pollutants* (Mosby, St Louis, 1973).
3. Webb, M. in *Topics in Environmental Health* Vol. 2 (Elsevier, Amsterdam, 1979).
4. Nriagu, J. O. in *Topics in Environmental Health* Vol. 1 (ed. Nriagu, J. O.) (Elsevier, Amsterdam, 1978).
5. Nriagu, J. O. in *Topics in Environmental Health* Vol. 3 (ed. Nriagu, J. O.) (Elsevier, Amsterdam, 1979).
6. Bull, K. R., Murton, R. K., Osborn, D., Ward, P. & Cheng, L. *Nature* **269**, 507–509 (1977).
7. Osborn, D., Harris, M. P. & Nicholson, J. K. *Comp. Biochem. Physiol.* **64C**, 61–67 (1979).
8. Nicholson, J. K., thesis, Univ. London (1980).
9. Nicholson, J. K. & Osborn, D. *J. Zool.* **202**, 99–118 (1983).
10. Osborn, D. in *The Management and Control of Heavy Metals in the Environment*, 187–190 (CEP Consultants, Edinburgh, 1979).
11. Osborn, D. *Biochem. Pharmac.* **27**, 822–824 (1978).
12. Nicholson, J. K. & Kendall, M. D. *J. appl. Tox.* **1**, xxiv (1981).
13. Kendall, M. D. & Nicholson, J. K. *J. Anat.* **132**, 480 (1982).
14. Kendall, M. D., Warley, A., Nicholson, J. K. & Appleton, T. C. *J. Cell Sci.* **62**, 319–338 (1983).
15. Banerjee, D., Onosake, S. & Cherian, M. G. *Toxicology* **24**, 95–105 (1982).
16. Nechay, B. R. & Saunders, J. P. *J. Pharmac. exp. Ther.* **200**, 623–629 (1977).
17. Aughey, E., Scott, R., Fell, G. S. & Cunningham, C. in *Trace Element Metabolism in Man and Animals* TEMA-4, 453–456 (Aust. Acad. Sci., 1981).
18. Friberg, L., Piscator, M., Nordberg, C. & Kjellstrom, T. *Cadmium in the Environment* (CRC, Cleveland, 1974).
19. Hamilton, E. I., Minski, M. J. & Clearly, J. J. *Sci. Total Envir.* **1**, 341–374 (1972).
20. World Health Organization *Tech. Rep. Ser.* 505 (WHO, Geneva, 1972).
21. Lauwerys, R. in *Topics in Environmental Health* Vol. 2 (Elsevier, Amsterdam, 1979).

Rapid photochemical inactivation of Ca^{2+} -antagonists shows that Ca^{2+} entry directly activates contraction in frog heart

M. Morad*, Y. E. Goldman* & D. R. Trentham†

Departments of Physiology* and Biochemistry & Biophysics†, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

'Calcium-antagonists' are a group of pharmacological agents which are potent vasodilators and are clinically used for the treatment of angina. They are thought to block Ca^{2+} channels in vascular smooth muscle^{1,2} and myocardium^{3–5} but other sites of action have been proposed⁶. These agents bind tightly to heart muscle and suppress action potential and contraction. Nifedipine and nisoldipine (BAY K 5552) are Ca^{2+} antagonists which have *o*-nitrobenzyl groups and are photolabile⁷. We have found that short pulses of UV light rapidly inactivate these drugs in ventricular muscle. This observation allowed us to

study the effect of Ca^{2+} antagonists on action potential, Ca^{2+} current and tension in conditions in which diffusion of those drugs from their site of action was not rate limiting. Our studies, described here, suggest that the primary mechanism of action of Ca^{2+} antagonists is the blockade of the Ca^{2+} channel and support the idea that extracellular space is the immediate source of contractile Ca^{2+} in the frog heart.

Frog ventricular muscle was used in the present study primarily because development of tension in this tissue is continuously controlled by the membrane potential, suggesting that intracellular Ca^{2+} pools do not contribute significantly to the development of tension^{8,9}. Frog ventricular strips (0.4–0.5 mm in diameter) were placed in a single sucrose gap chamber⁸ and trans-gap potential and developed tension were continuously monitored. In some experiments the membrane potential was clamped using a chopped current voltage-clamp system¹⁰. The light source for photoconversion of the applied drugs was either a shuttered 200-W continuous mercury arc lamp or a 5-ms xenon flash. IR radiation was filtered with a 10-cm water filter and in some experiments the light also passed through a 300–400-nm glass filter.

Figure 1 shows twitch tension recordings from a strip of frog ventricular muscle which was stimulated continuously at a frequency of 12 shocks per min in Ringer solution. At the time indicated on the tension records, 1 μM nifedipine was added to the external medium, suppressing the twitch tension slowly. At the second arrow the muscle was exposed to a 1-s pulse of 300–400 nm light. Twitch tension recovered in a step-like fashion, even though the suppression of tension by the drug required about 10 min. If the drug was removed from the bathing medium without a light flash, recovery of tension required 30–40 min. Irradiation during the washout period caused immediate and often complete recovery of tension. This finding suggests that drug molecules bound to their active site are inactivated by irradiation. Pre-irradiated nifedipine solution had no effect on the twitch tension. In untreated preparations, similar pulses of light often enhanced the action potential and contraction. This effect was small and variable (5–10% increment) and disappeared after several flashes. The suppression of tension by nifedipine and its removal by short intense light pulses occurred whether or not light initially had a direct inotropic effect on a particular muscle.

Nifedipine and nisoldipine each contain *o*-nitrobenzyl groups. These groups are known to undergo intramolecular photochemical redox reactions that lead to overall photolysis or photoisomerization¹¹. Irradiation of nifedipine or nisoldipine results in substituted pyridine and loss of water with a concomitant spectral change⁷ (Fig. 1). The kinetics of the photochemical reactions, at 22°C in water buffered at pH 7, were measured from the change of absorption at 400 nm following irradiation of nifedipine or nisoldipine with a 50-ns pulse from a frequency-doubled ruby laser. The spectral change was complete within 100 μs (the time resolution of the spectrophotometer). The photochemical reaction which leads to aromatization of the dihydropyridine group is much more rapid than the corresponding reaction that occurs when *o*-nitrobenzyl groups are photolysed to nitrosoketones^{12,13}. Partition experiments between water and chloroform suggested that nifedipine and its photoproduct are soluble in the membrane. Nevertheless, the tension recordings of Fig. 1 show that the photoproduct of bound nifedipine did not suppress twitch tension.

Nimodipine (BAY E9736), another Ca^{2+} antagonist but with a *m*-nitrobenzyl group, did not show an absorption change on irradiation. Photoinactivation of nimodipine bound to ventricular muscle was possible but required at least 10-fold longer pulses of light for recovery of twitch tension. Diltiazem, a benzothiazepine type of Ca^{2+} antagonist, showed no light sensitivity.

The effectiveness of light-induced removal of the suppressive effect of nifedipine or nisoldipine on twitch tension was examined at various times in the cardiac cycle by using brief pulses of light. In Fig. 2, each panel shows a sequence of three

Fig. 1 Photoconversion of nifedipine. Tension records show twitches of an electrically stimulated (12 per min) frog ventricular strip bathed in Ringer solution containing (mM) Na^+ , 115; Cl^- , 118; K^+ , 3; HCO_3^- , 2; Ca^{2+} , 1; pH 7.1 at 20 °C. Nifedipine (Nif), 1 μM , was added at the time indicated by the arrow. Twitch C (also shown in expanded time scale) is the suppressed twitch before exposure of the muscle to a 1-s pulse of 300–400 nm light. Twitch f is the first beat after the light pulse. In this experiment the bathing solution was not flowing. The photochemical reaction of nifedipine is shown together with the accompanying UV spectral change recorded in 100 mM TES buffer at pH 7.0. Photolysis of nisoldipine

(which has an identical structure to nifedipine except that one of the methyl esters is replaced by an isobutyl ester) was also accompanied by a similar UV spectral change. The photoproduct of nisoldipine was crystallized and had an identical UV spectrum to the photoproduct in the mother liquor, indicating that the photolysis results in a single product (plus water). The structure of the nisoldipine photoproduct was confirmed by proton NMR measured in deuterated chloroform. In particular, single protons 4.14 (NH) and 4.20 (both singlets) in the nisoldipine spectrum were absent in that of the photoproduct.

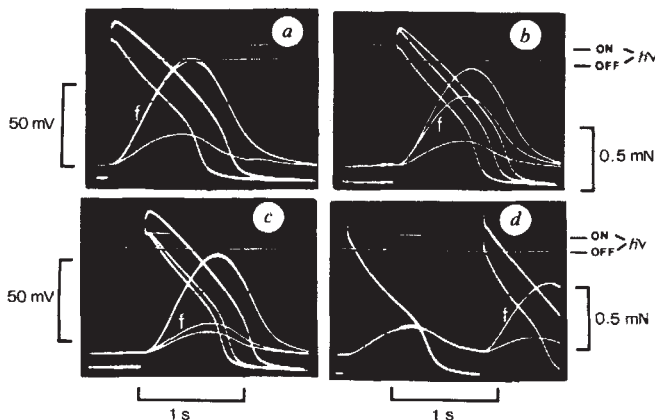
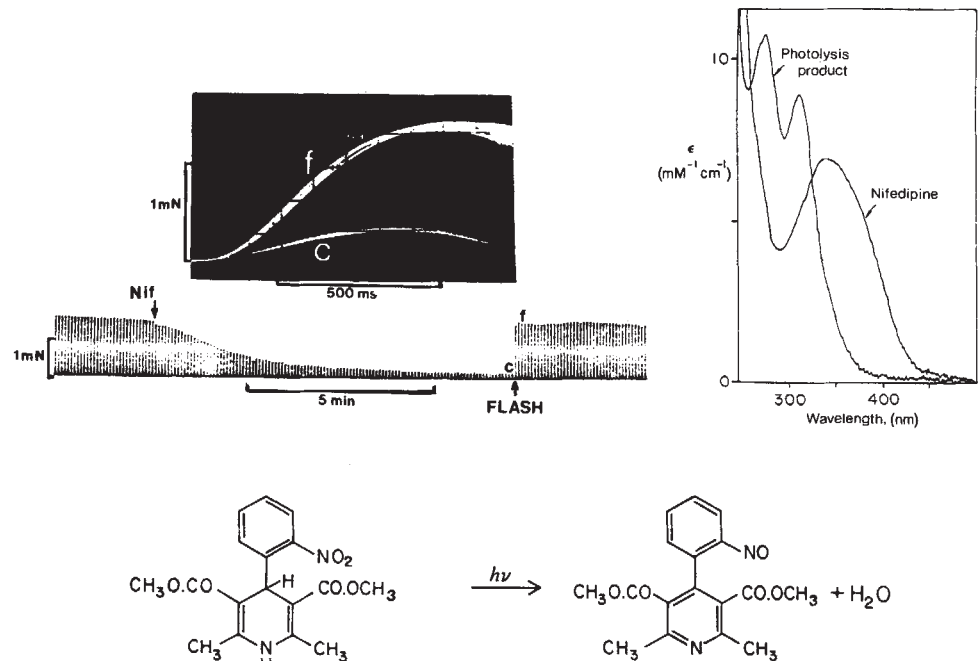


Fig. 2 Kinetics of photoconversion of nifedipine. In each panel upper traces show action potentials measured across a single sucrose gap and lower traces show the developed tension. The horizontal bars indicate 200-ms light pulses from a 200-W mercury arc lamp. *a–c* Each show three superimposed action potentials and twitches of consecutive beats occurring at 5-s intervals. Lowest action potential and contraction are the steady-state suppression in 1 μM nifedipine. Twitch labelled f is the beat stimulated immediately after (b), simultaneous with (c) or 4 s after the light pulse (a). The highest contraction and action potential occurred in each case 5 s after the f beat. These beats are not fully recovered because continuous perfusion of non-irradiated drug maintains a partial suppressive effect. *d* Shows paired pulse stimulation occurring at 5-s intervals. The lower pair indicates the steady-state suppression. The 200-ms pulse of light was triggered during the first action potential of the second pair. Trace f is the recovered second beat of that pair. Other conditions are as in Fig. 1.

successive action potentials and contractions during a photoactivation procedure. In Fig. 2a, a 200-ms pulse of light (horizontal bar) was triggered before the action potential at $t = -4$ s. All of the recovery occurred in the first beat (f) after the light pulse. In Fig. 2b, the pulse of light was applied just before the twitch at $t = -0.2$ s. Most of the recovery occurred with the first beat following the illumination. If, on the other hand, the light pulse was applied during the action potential ($t = 0$, Fig. 2c), little recovery from the nifedipine block was observed in that beat (f). However, the subsequent beat, elicited 5 s later, was fully recovered (largest twitch). Paired pulsing of the preparation (Fig. 2d) showed that a brief period of repolarization was required for unblocking to occur. Full mechanical relaxation was not necessary for recovery.

The interval during diastole required for recovery of tension after a pulse of light was examined for nisoldipine at higher time resolution with a 5-ms xenon flash lamp. The results plotted in Fig. 3 indicate that inactivation of nisoldipine is equally effective at any time during diastole until about 100 ms before the stimulus (Fig. 3). Irradiation during the action potential had little effect on the accompanying twitch. However, the twitch of the subsequent beat recovered to the same extent as

when the flash was applied in diastole. The graph indicates that on depolarization of the membrane an abrupt transition occurs in the extent of light-induced recovery of tension. Results of Figs 2 and 3 show that illumination is equally effective in photoconverting nifedipine or nisoldipine at rest or during the action potential. However, molecules of nifedipine or nisoldipine which have absorbed a UV photon apparently continue to suppress excitation-contraction coupling until the membrane repolarizes.

It might be argued that the reason for the reduction of immediate recovery when light pulses were applied during the action potential is that the Ca^{2+} -transport mechanism may have been inactivated during the depolarization. This possibility was considered unlikely as it has already been shown that development of tension responds rapidly and continuously to changes in membrane potential throughout very long (5–10 s) depolarizing pulses in this tissue^{8,9}.

The photoactivation technique allowed us to test whether nifedipine or nisoldipine operate on the surface membrane or intracellularly by rapidly inactivating them in the presence of Ni^{2+} , a membrane-impermeable Ca^{2+} -channel blocker^{14–16}. The advantage of using Ni^{2+} (1–2 mM) compared with other inor-

Fig. 3 Kinetics of twitch recovery after photoinactivation of nisoldipine. A 5-ms pulse from a xenon flash lamp was used to photoinactivate nisoldipine in a preparation stimulated at 12 shocks per min. The inset shows superimposed action potentials and twitches for a light pulse at the time indicated by f. T_1 is the steady-state suppressed twitch tension. T_2 is the twitch tension of the first beat after the light pulse. T_3 is the twitch tension of the next beat stimulated 5 s later. Circles: the per cent T_2 recovery, $100(T_2 - T_1)/(T_3 - T_1)$, is plotted against the time between the stimulus which elicited T_2 and the light pulse. Data plotted to the right of the ordinate were derived from the twitches accompanying the light pulse. Zero time and $t = -5$ s correspond to the upstroke of the regularly stimulated action potentials. Each point to the right of the abscissa has a corresponding recovery of the next beat plotted to the left of the ordinate as squares. As the 5-ms flash did not produce as much energy as the 0.2–1.0-s shuttered flash, T_3 does not represent full recovery but about 70% of the control tension before nisoldipine was applied. Other conditions as in Fig. 2.

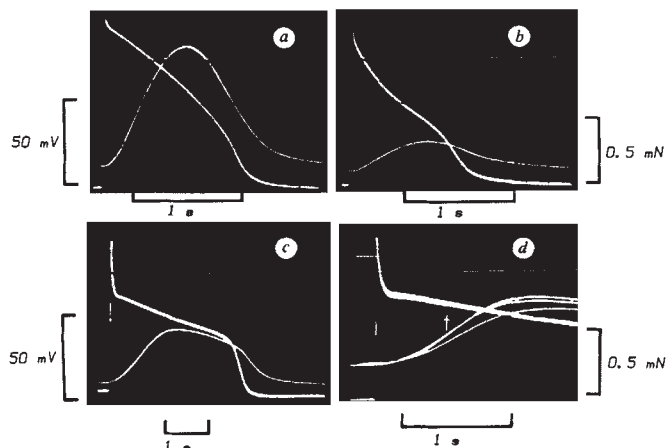
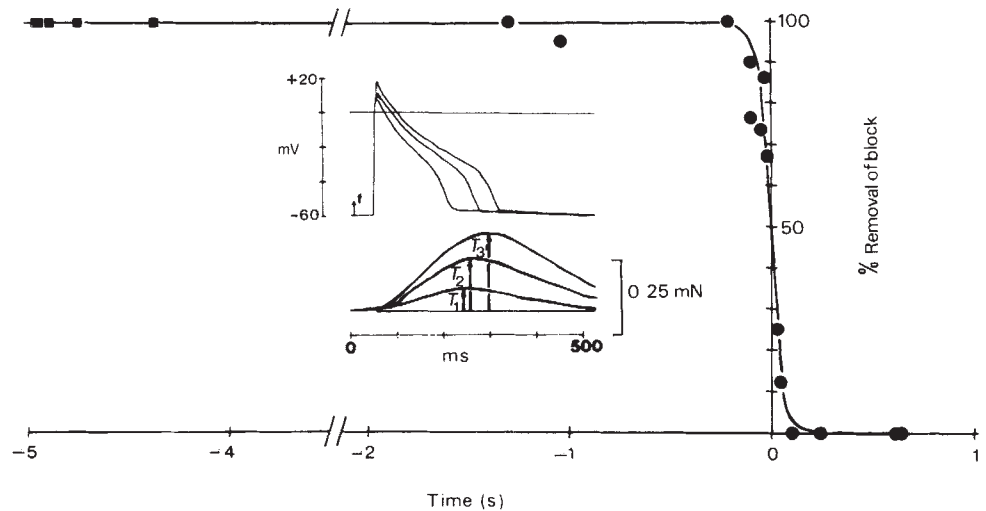


Fig. 4 Interaction of nifedipine and Ni^{2+} . In *a*, trans-gap action potential (upper traces) and the accompanying contraction (lower trace) are indicated. *b* Shows the steady-state suppression of $1 \mu\text{M}$ nifedipine on both action potential and contraction. In *c* the preparation has been exposed for 30 min to Ringer's containing $1 \mu\text{M}$ nifedipine and 2 mM Ni^{2+} . Action potential is prolonged three- to fourfold, and although the initial rate of development of tension is suppressed (compare *d*, lower trace, and *b*), the peak tension is larger than in the presence of nifedipine (*b*). In *d* (expanded time base) three successive action potentials and contractions are superimposed in the presence of nifedipine and Ni^{2+} . The first, before the flash (control), was accompanied by the smallest twitch. The second and third, occurring after a 200-ms light pulse (horizontal bar), were accompanied by contractions with slightly higher initial rates of rise of twitch tension (*f*). For comparison of the effect of similar light flashes in the absence of Ni^{2+} , see Fig. 2*a* which involves the same preparation.

ganic Ca^{2+} -channel blockers is that the blockade of Ca^{2+} current in Ni^{2+} -treated ventricular strips is accompanied by prolongation of the action potential (3–4 s). Twitches accompanying such action potentials have much lower rates of rise (dp/dt) but generate considerable tension^{14,17}. When Ni^{2+} was added to nifedipine-treated preparations, dp/dt decreased, the action potential was prolonged and the peak tension remained the same or increased slightly (compare Fig. 4*c* and *d* with *b*). Irradiation of the muscle bathed in Ni^{2+} and nifedipine-containing solutions caused no appreciable change in the action potential, although a very slight increase in the rate of rise of twitch (*f*) was often seen (Fig. 4*d*, compare with Fig. 2*a*, same

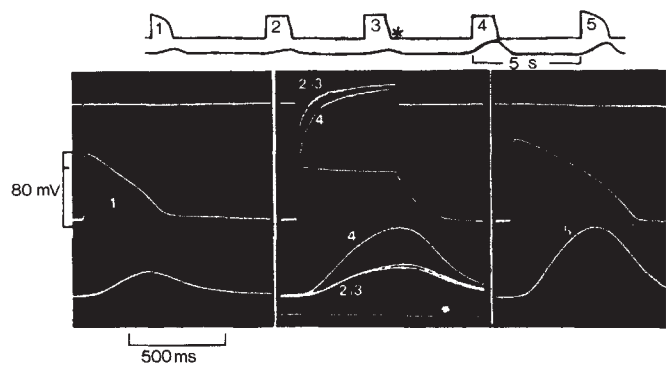


Fig. 5 Recovery of inward current by photoinactivation of nisoldipine. Upper diagram shows the protocol of action potential and clamp pulses applied at 5-s intervals to a ventricular muscle equilibrated in $1 \mu\text{M}$ nisoldipine. The control (nisoldipine-suppressed) action potential and contraction are shown as traces 1 (left panel, original recordings). The sequence of stimulation is then interrupted by a series of three clamp pulses before resumption of regular stimulation (see upper diagram). After the second clamp pulse the muscle is exposed to a 200-ms light pulse (horizontal bar with *). Superimposed traces of membrane current, voltage and developed tension are shown in the middle panel and are labelled 2, 3 and 4 in correspondence with the diagram. Note that before the flash, tensions and membrane currents superimpose for clamp pulses 2 and 3 to 0 mV. However, after the flash, membrane current is more inward and developed tension is enhanced for a clamp pulse to the same potential (traces 4). The right panel shows the enhanced action potential and tension of the next normally stimulated beat (traces 5). Experimental conditions are the same as those of Fig. 2.

preparation in nifedipine but without Ni^{2+}). Similar effects were also observed with nisoldipine or if Ni^{2+} was added before the Ca^{2+} antagonists. These results suggest that Ca^{2+} antagonists suppress tension primarily by blocking the Ca^{2+} channel in the surface membrane, because otherwise their rapid photoinactivation in the presence of Ni^{2+} would have resulted in substantial recovery of tension.

To confirm that the slow calcium current was involved in photochemical reversal of drug suppression, several preparations were voltage clamped. A 200-ms flash of UV light, applied before the clamp, caused a marked enhancement in the inward current (Fig. 5, beat 4). Thus, the enhancement in the

duration and the amplitude of the action potential and contraction seem to be related to an increase in a slowly inactivating inward current. Similar results were obtained in the presence of tetrodotoxin, which inhibited the fast Na^+ current. We conclude, therefore, that photoconversion of nifedipine or nisoldipine results in release of the block on the slow inward current which, in turn, causes the recovery of action potential and contraction. This conclusion is supported by studies in which organic Ca^{2+} antagonists were found to have little or no effect on the development of tension in detergent- or EGTA-skinned myocardium¹⁸⁻²⁰.

Rapid photochemical inactivation of Ca^{2+} antagonists provides a unique tool for studying the characteristics and kinetics of the Ca^{2+} channel in excitable tissues. Unblocking of the drug-sensitive site seems to be markedly influenced by the membrane potential. This observation suggests a direct voltage-

dependence of the dissociation of the photo-product to the bulk phase of the membrane or to adjacent aqueous compartments. It is possible also that the channel opening or channel currents may affect unblocking.

The immediate and essentially full recovery of tension within one beat after a short flash of light (Fig. 1) suggests that the drug blocks the pathway of Ca^{2+} transport which directly activates the myofilaments. The voltage-clamp experiments and the results using Ni^{2+} (Figs 4, 5) indicate that the blocked-transport system is the Ca^{2+} channel in the surface membrane and that the most likely source of activator Ca^{2+} in the frog heart is the extracellular space.

We thank Miss Joan Thomas, Mr Taylor Allen, also Dr A. Scriabine and Miles Laboratory for nisoldipine and nimodipine. This work was supported by NIH grants HL16152, AM00745, HL15835 and Whitehall Foundation.

Received 17 February; accepted 30 June 1983.

1. Bolton, T. B. *Physiol. Rev.* **59**, 606-718 (1979).
2. Meisheri, K. D., Hwang, O. & Van Breeman, C. J. *Membrane Biol.* **59**, 19-25 (1981).
3. Coraboeuf, E. *The Slow Inward Current and Cardiac Arrhythmias*, Ch. 3 (Martinus Nijhoff, Boston, 1980).
4. Henry, P. D. *Am J. Cardiol.* **46**, 1047-1058 (1980).
5. Fleckenstein, A. *New Perspectives on Calcium Antagonists*, Ch. 5 (Williams & Wilkins, Baltimore, 1981).
6. Millard, R. W. *et al.* *Am. J. Cardiol.* **49**, 499-506 (1982).
7. Schlossmann, V. K. *Arzneim.-Forsch. (Drug Res.)* **22**, 60-62 (1972).
8. Morad, M. & Orkand, R. J. *Physiol., Lond.* **219**, 167-189 (1971).
9. Morad, M. & Goldman, Y. *Prog. Biophys. molec. Biol.* **27**, 257-313 (1973).
10. Goldman, Y. & Morad, M. J. *Physiol., Lond.* **268**, 613-654 (1977).
11. Morrison, H. A. *The Chemistry of the Nitro and Nitroso Groups Pt 1* (ed. Feuer, H.) 165-213 (Wiley, New York, 1969).
12. Kaplan, J. H., Forbush, B. III & Hoffman, J. F. *Biochemistry* **17**, 1929-1935 (1978).
13. McCray, J. A., Herbette, L., Kihara, T. & Trentham, D. R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7237-7241 (1980).
14. Morad, M. & Klitzner, T. *Biophysical Aspects of Cardiac Muscle* (ed. Morad, M.) 1227-1243 (Academic, New York, 1978).
15. Kohlhardt, M., Bauer, B., Krause, A. & Fleckenstein, A. *Pflügers Arch. ges. Physiol.* **338**, 115-123 (1973).
16. Kohlhardt, M., Mnich, Z. & Haap, K. J. *molec. cell. Cardiol.* **11**, 1227-1243 (1979).
17. Babskii, E. B. & Donskikh, E. A. *Doklady Akademii Naak. SSSR* **178**, 248-251 (1968).
18. Morad, M., Tung, L. & Greenspan, A. M. *Am. J. Cardiol.* **49**, 595-601 (1982).
19. Ishizuka, T. & Endo, M. *Circulation Res.* **52**, Suppl. 1, 110-114 (1983).
20. Nayler, W. G. & Grinwald, P. *Fedn Proc.* **40**, 2855-2861 (1981).

Identification and probable role of a single neurone containing the neuropeptide *Helix* FMRFamide

G. A. Cottrell*, L. P. C. Schott† & G. J. Dockray‡

* Department of Physiology and Pharmacology, University of St Andrews, Fife KY16 9TS, UK

† Biologisch Laboratorium, Vrije Universiteit de Boelelaan, Amsterdam, The Netherlands

‡ The Physiological Laboratory, University of Liverpool, Liverpool L69 3BX, UK

Recently, there has been intense interest in the possibility that peptides might function as neurotransmitters^{1,2}. Despite much progress, there remains no clear-cut example in which the production of a chemically characterized peptide may be ascribed to individual identifiable neurones of proven physiological role. Invertebrate systems have proved to be particularly valuable for the study of identified neurones, including those producing peptides^{3,4}. We have now identified a neurone in ganglia of *Helix* that is associated with a peptide of the Phe-Met-Arg-Phe-NH₂ (FMRFamide) group, and which influences tentacle contraction. This system offers for the first time the capacity to study peptidergic transmission in a system in which both the cell soma and the postsynaptic target may be readily and reproducibly identified.

The molluscan neuropeptide FMRFamide, isolated from the clam *Macrocallista nimbosa*⁵, has potent effects on smooth muscle⁶, cardiac muscle⁷ and neurones (ref. 8 and G.A.C. and N. Davies, unpublished). Immunocytochemical studies in the snails *Lymnaea stagnalis*⁹ and *Helix aspersa*¹⁰ provide clear evidence for a neuronal localization of a FMRFamide-like peptide in these species. The *Helix* peptide has a different structure from that of *Macrocallista* (see ref. 11) and Price¹² has recently suggested that it is pGlu-Asn-Phe-Ile-Arg-Phe-NH₂.

In the present study, immunocytochemical experiments were done as described before, using antisera to FMRFamide¹³. Sections of central ganglia were carefully examined to locate single large neurones which could possibly be identified in intact, unfixed preparations. One neurone which reacted

strongly with the antibody was observed on the dorsal surface of each cerebral ganglion. The neurone (C3) was positively identified in living isolated ganglion preparations by injecting Procion yellow and subsequently processing the ganglia for FMRFamide immunoreactivity (Fig. 1a,b). An electron microscopic study showed that the C3 neurone contains elementary granules with a mean diameter of 100 nM (Fig. 1c).

Confirmation of the peptide content of this neurone was obtained by radioimmunoassay using antibody specific for the C-terminus of FMRFamide that reacts with the *Helix* peptide^{14,15}. Individual C3 neurones were isolated by dissection with fine forceps and sharpened fine needles and placed in 1 ml of distilled water, heated in a boiling water bath for 3 min and acidified by addition of 10 μ l of acetic acid. In six cells the mean content was 367 $\pm$ 90 fmol (mean $\pm$ s.e.). By comparison, in two neurones that were not stained by immunohistochemistry, C1 (see ref. 16) and F1 (ref. 17), the FMRFamide-immunoreactive content was 54 $\pm$ 12 and 90 $\pm$ 21 fmol respectively, that is, 4-10 times less than for C3. The processes of the left and right C3 neurones were traced within the ganglia and associated nerve trunks by injecting Lucifer yellow into the perikaryon, fixing in buffered (pH 7.4) 4% glutaraldehyde, dehydrating in an ethanol series and clearing and mounting in methyl salicylate. Each neurone appeared to be symmetrical, mirror-image of the one on the other side. Figure 2 shows a drawing obtained from a dye-filled C3 neurone located in the left cerebral ganglion. The features of immediate interest are that (1) the main process circles within the central part of the ganglion and (2) it sends several branches along a nerve to the tentacle muscle. Most of the processes pass into the tentacle muscle, where they become very faint and difficult to follow. Some of the processes may also innervate associated small muscle bands close to the point of junction of the nerve trunk to the tentacle muscle. These processes are about 7 mm from the perikaryon of C3. Immunocytochemical experiments provide clear evidence for the localization of FMRFamide immunoreactivity in fine neuronal processes within the tentacle muscle (Fig. 1d). The proportion of such fine immunoreactive processes within the muscle which originate from the C3 neurone is, however, unknown.

Cerebral ganglia were isolated together with attached tentacle muscles and the entire preparation was immersed in a continuous flow of physiological solution (ref. 18) in a small

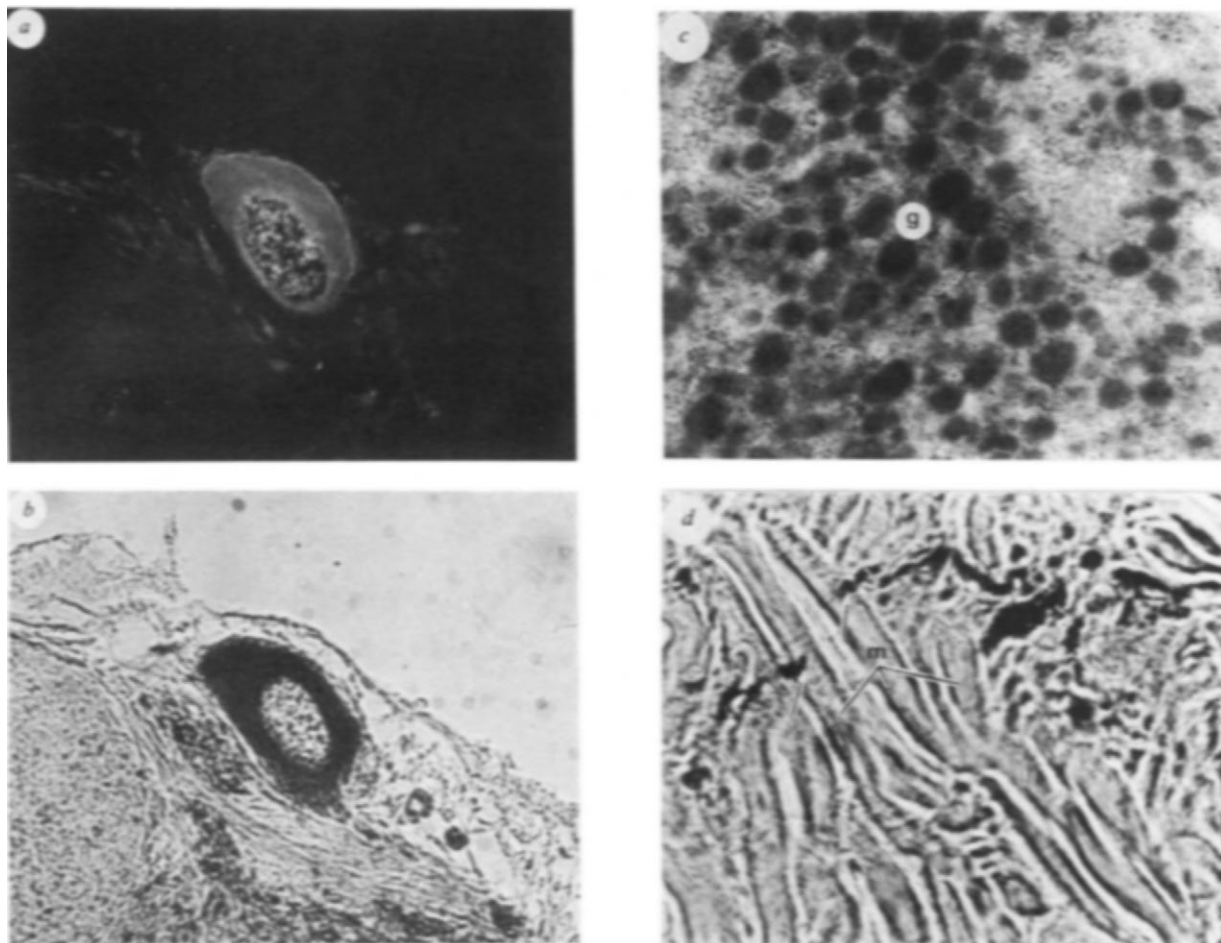


Fig. 1 *a, b*, Serial sections of part of a cerebral ganglion which contains the C3 neurone. *a* Was photographed under dark background UV illumination and shows the presence of Procion yellow, which was injected previously from a micropipette. *b*, Appearance of the same neurone in an adjacent section which has been processed for the immunocytochemical localization of FMRFamide. The dark staining is indicative of a positive reaction. *c*, Electron micrograph of the cytoplasm of the C3 neurone showing many peptidergic elementary granules (g) ($\times 42,500$). *d*, Anti-FMRFamide reactive neuronal processes lying among muscle fibres of the tentacle muscle. Some processes seem to end, probably synaptically, on muscle fibres (arrows) ($\times 370$).

Perspex chamber. Tension of the tentacle muscle was monitored by means of a Grass force-displacement transducer. One, or both, of the two symmetrically located C3 neurone perikarya was impaled with a 1 M potassium acetate micropipette. Stimulation of one of the neurones resulted in contraction of the ipsilateral tentacle muscle alone. No evidence was obtained for a connection between the two C3 neurones. The rate, and to some extent the degree, of contraction of the ipsilateral muscle was dependent on the frequency of impulse activity in the appropriate C3 neurone (see Fig. 3). One interesting feature of the electrical activity recorded from the C3 neurone is the intense synaptic innervation they receive.

The effects of FMRFamide and some other pharmacologically active compounds have been tested on the mechanical activity of the isolated tentacle muscle. FMRFamide in low doses (≥ 1 nM) caused the muscle to contract (Fig. 4*a*), whereas [Met]enkephalin, for instance, even in doses of 10 μ M, was without effect and serotonin at doses above 10 nM caused relaxation. SCP_B, a cardioactive peptide (see ref. 19), in micromolar concentrations had a slight relaxing effect on the muscle. Sometimes, during prolonged exposure, the response to FMRFamide consisted of a prolonged contraction on which were superimposed rhythmical contractions. Acetylcholine (ACh) also contracted the muscle, but was ~ 100 times less potent than FMRFamide. The ACh antagonist benzoquinonium chloride (Mytolon; Sterling-Winthrop) potentially blocked the action of ACh, but did not affect the response of the muscle to FMRFamide (not shown). Benzoquinonium chloride, even in doses up to 100 μ M, did not block the response of the muscle to activation of the C3 neurone (Fig. 4*b*).

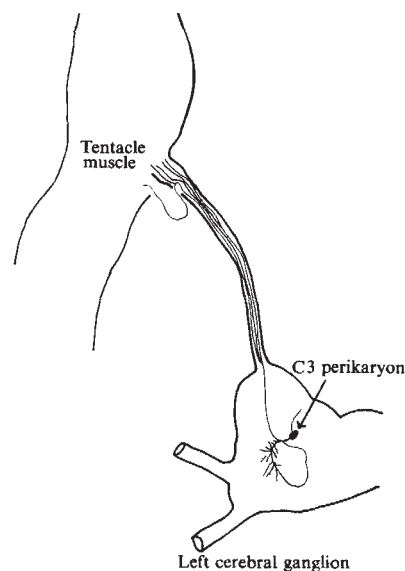


Fig. 2 A tracing of the processes of the C3 neurone of the left cerebral ganglion after injection with Lucifer yellow. After making a small loop in the neuropile of the ganglion, one process passes into a nerve trunk and branches. These branches pass towards the tentacle muscle and many of them pass into the muscle. Another process passes antero-medially and appears to end abruptly. This may be because it passes into a small nerve, or into the connective tissue capsule, which has been removed to allow insertion of the microelectrode.

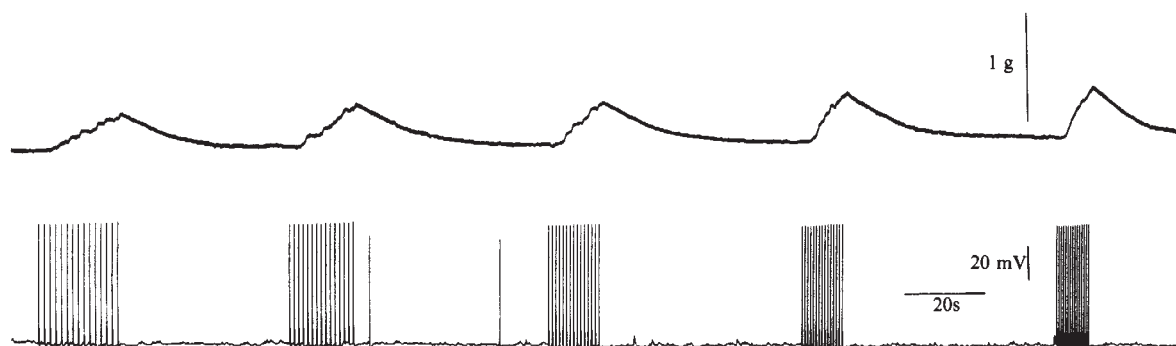


Fig. 3 Influence of impulse activity in the C3 neurone of a left cerebral ganglion on the mechanical activity of the left tentacle muscle. Impulse activity in the neurone (lower recording), produced by injecting brief depolarizing pulses, caused the muscle to contract. Increasing the impulse frequency from 1.6 Hz to 3.1 Hz increased the rate of contraction and also, slightly, the strength of contraction.

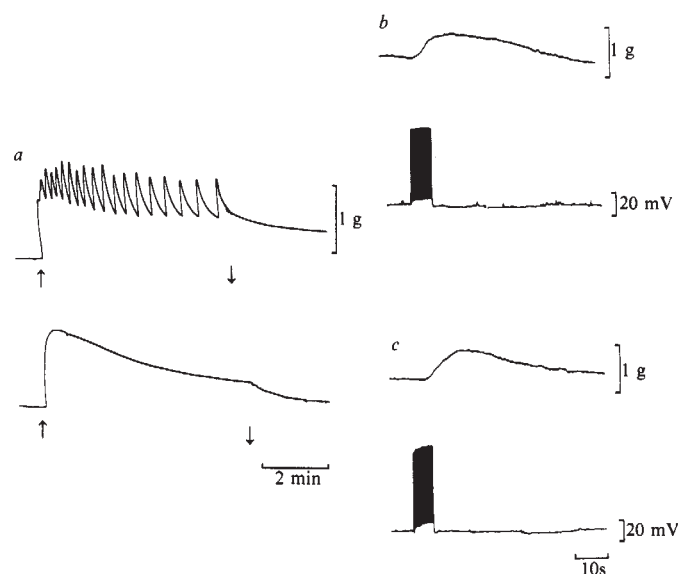


Fig. 4 *a*, Upper recording shows the response of an isolated tentacle muscle to 80 nM FMRamide. Tension was recorded isotonicity and the preparation immersed in Meng's saline¹⁸. During long-term exposure, rhythmic contractions were often seen superimposed on a background increase in tension, as shown here (note time scale). Lower recording, response of the muscle to 5 μM ACh. After the initial increase in tension, there followed a gradual decline in tension before the ACh was washed from the bath. *b*, Upper recording shows the response of the left tentacle muscle to a burst of spikes in the left C3 neurone (lower trace) in normal physiological solution. *c*, These recordings should be compared with those in *b*. They were made after exposure to benzoquinonium chloride at 100 μM for 16 min. At this concentration of benzoquinonium chloride the response to ACh, even at 1 mM, was reduced to less than 15%. The drug did not, however, antagonize the effect of the C3 neurone input.

The results presented therefore suggest the existence of a simple system, comprising two symmetrically located neurones and the muscle which each activates, that is suitable for detailed study of the role of the *Helix*-FMRamide-related peptide. Available data suggest, moreover, that one role of this neuropeptide is that of a motoneurone 'transmitter substance'.

We thank the MRC and the European Science Foundation for financial support, Professor H. H. Boer for useful discussions and Mr A. Oates for some of the data on the effect of benzoquinonium chloride on the response of the muscle to ACh.

Received 27 April; accepted 29 June 1983.

- Hokfelt, T., Johansson, O., Ljungdahl, A., Lundberg, J. M. & Schultzberg, M. *Nature* **284**, 515-521 (1980).
- Snyder, S. H. *Science* **209**, 976-983 (1980).

- Osborne, N. N. (ed.) *Biochemistry of Characterised Neurones* (Pergamon, Oxford, 1980).
- O'Shea, M. & Adams, M. E. *Science* **213**, 567-569 (1981).
- Price, D. A. & Greenberg, M. J. *Science* **197**, 670-671 (1977).
- Price, D. A. & Greenberg, M. J. *Gen. Pharmac.* **11**, 237-241 (1980).
- Painter, S. D. & Greenberg, M. J. *Biol. Bull.* **162**, 311-332 (1982).
- Cottrell, G. A. *Nature* **296**, 87-89 (1982).
- Boer, H. H., Schot, L. P. C., Veenstra, J. A. & Reichelt, D. *Cell Tissue Res.* **213**, 21-27 (1980).
- Marchand, C.-R., Wijdenes, J. & Schot, L. P. C. *C. r. heb. Séanc. Acad. Sci., Paris* **294**, 34-44 (1982).
- Cottrell, G. A., Price, D. A. & Greenberg, M. J. *Comp. Biochem. Physiol.* **70C**, 103-107 (1981).
- Price, D. A. *Comp. Biochem. Physiol.* **72C**, 325-328 (1982).
- Schot, L. P. C., Boer, H. H., Swaab, D. F. & van Noorden, S. *Cell Tissue Res.* **216**, 273-291 (1981).
- Dockray, G. J., Vaillant, C. & Williams, R. G. *Nature* **293**, 656-657 (1981).
- Dockray, G. J., Vaillant, C., Williams, R. G., Gayton, R. J. & Osborne, N. N. *Peptides* **2**, Suppl. 2, 25-30 (1981).
- Cottrell, G. A. & Osborne, N. N. *Nature* **225**, 470-472 (1970).
- Kerkut, G. A., Lambert, J. D. C., Gayton, R. J., Locker, J. E. & Walker, R. J. *Comp. Biochem. Physiol.* **50A**, 1-25 (1975).
- Meng, K. *Zool. Jb.* **68**, 539-566 (1960).
- Morris, H. R., Panico, M., Karplus, A., Lloyd, P. E. & Riniker, B. *Nature* **300**, 643-645 (1982).

Long-term regenerated nerve fibres retain sensitivity to potassium channel blocking agents

Jeffery D. Kocsis & Stephen G. Waxman

Department of Neurology, Stanford University School of Medicine and Veterans Administration Medical Center, Palo Alto, California 94304, USA

Mammalian myelinated peripheral nerve fibres display a remarkable degree of regeneration following a discrete nerve crush¹. Nerve crush disrupts the axon cylinder, but leaves the basement membrane of the Schwann cell intact. These intact endoneurial tubes provide pathways to guide the regenerating axon sprouts. After contact with the periphery is established, the regenerating fibres enlarge and myelinate. Conduction velocity recovers to nearly normal and functional recovery is, in many cases, nearly complete. A distinct feature of normal mature myelinated axons is the insensitivity of these fibres to potassium channel blocking agents²⁻⁴. In contrast, immature myelinated axons are exquisitely sensitive to the K channel blocking agent 4-aminopyridine (4-AP)^{4,5}. Application of 4-AP to immature myelinated fibres leads to a delayed membrane depolarization with action potential burst activity in response to a single stimulus⁵. This sensitivity to 4-AP is attenuated as the fibres mature. Previous studies^{4,6} have demonstrated a sensitivity to 4-AP in regenerating nerve fibres; this sensitivity differentiates the regenerating axon segments from their normal parent axon segments. Such studies have not, however, examined the question of whether regenerated fibres, which have re-established peripheral connections and are functionally active, fully recapitulate the functional organization of normal

mature myelinated fibres. We demonstrate here that while sensitivity to the potassium channel blocking agents 4-AP and 3, 4-diaminopyridine (3, 4-DAP) is lost in the normal course of myelinated axon maturation, this property is present in long-term regenerated axons. This suggests that long-term regenerated mammalian axons are characterized by a functional organization that bears a closer resemblance to that of immature myelinated fibres than to that of adult myelinated fibres.

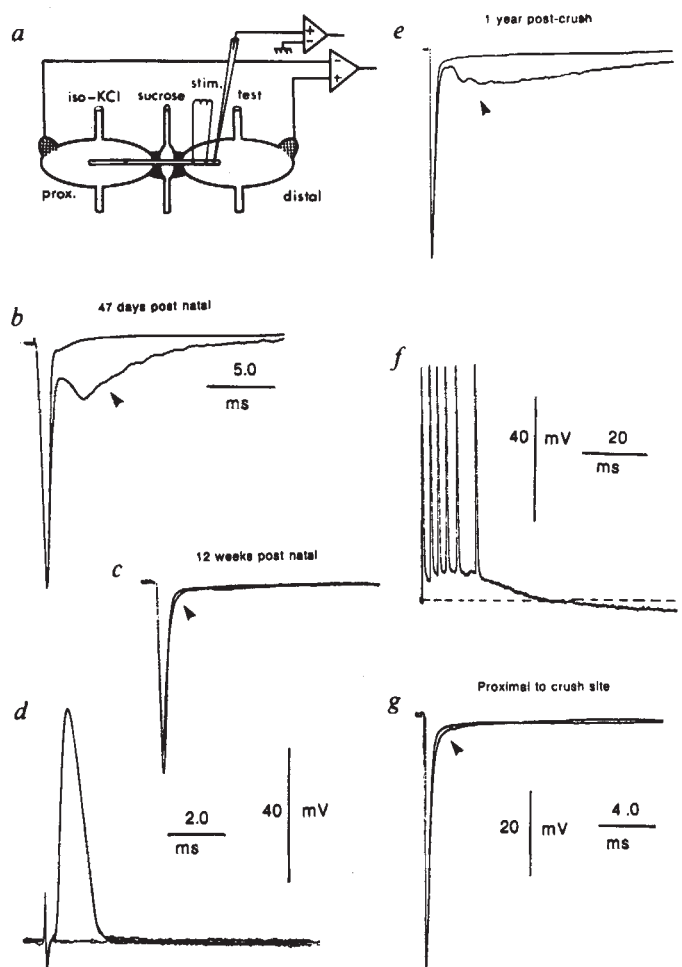
When 4-AP is applied to myelinated fibres of normal young animals, there is a significant change in the waveform of the compound action potential elicited by a single whole nerve stimulus. This is evident in Fig. 1*b* where a pronounced late negativity (arrowhead) emerges after 4-AP application (0.5 mM) to the sciatic nerve of a 47-day-old rat. The late negativity induced by 4-AP often gives the compound action potential a 'rippled' appearance, and from intra-axonal analysis has been shown to represent the extracellular correlate of bursts of action potentials elicited by a single stimulus^{5,6}. Application of the aminopyridine analogue 3,4-DAP gave rise to burst activity in a similar manner to 4-AP. These effects diminish during the course of maturation. In contrast, 4-AP application (1.0 mM) has very little influence on the waveform of the

compound action potential or on the waveform of the intra-axonally recorded action potential of normal adult mammalian myelinated nerve fibres⁶. The superimposed traces of compound action potentials recorded before and after (arrow) 4-AP application to a mature (12-week-old) nerve (Fig. 1*c*) demonstrate that 4-AP has minimal effect on the waveform of the response. There is no broadening of the initial negativity of the response, and only a slight increase in amplitude of the late activity. The intra-axonally recorded action potential in Fig 1*d* was recorded from the nerve of a mature (13-week-old) rat after 30 min exposure to 4-AP. There is no recognizable difference in the waveform of this action potential from that of others recorded in normal Ringer solution without 4-AP⁶, and only a single spike is elicited following a single stimulus.

In contrast to mature fibres, the action potentials of regenerated myelinated axons show a pronounced change in waveform after 4-AP application. In Fig. 1*e*, the arrow indicates the response recorded after 4-AP application to a nerve that was allowed to regenerate for slightly more than 1 yr. The pattern of the emergent late activity is similar to that observed in young animals (Fig. 1*b*), that is, it is characterized by the development of repetitive late activity in response to a single stimulus. During

Fig. 1 *a*, Diagram showing recording arrangement. The proximal (prox.) ends of the nerves were placed in isotonic KCl (iso-KCl) and the distal ends were placed in the test solution (compartment on the right). The central region was washed with isotonic sucrose. The three compartments were separated by petroleum jelly. All solutions ran continuously at a rate of 1–2 ml min⁻¹. The nerve was stimulated in the test compartment and intra-axonal recordings were obtained from the same compartment. Compound action potentials and d.c. potentials were recorded between the outer compartments through a high input impedance electrometer. *b*, Compound action potentials recorded before and after 4-AP (0.5 mM; arrow) application of a sciatic nerve from an immature rat (47 days postnatal). *c*, Recordings before and after (arrow) 4-AP application (1.0 mM) to the sciatic of a 12-week-old rat. *d*, Intra-axonal action potential recordings obtained from sciatic nerve of mature (13 weeks) rat in the presence of 4-AP. Stimulus intensity was adjusted to threshold; supra- and subthreshold responses are shown. *e*, Compound action potentials recorded before and after (arrow) 4-AP application to a sciatic nerve segment where regeneration has taken place for 1 yr. Note the emergence of the late activity after 4-AP application. *f*, Intra-axonal action potential burst recorded from regenerated nerve (1 yr post-crush) in the presence of 4-AP. Note the depolarizing potential from which the burst arises. *g*, Compound action potentials recorded before and after (arrow) 4-AP application to a nerve segment proximal to the crush site of a nerve crushed 1 yr previously. Time calibration in *b* pertains to *b* and *c*. Calibrations in *d* are for *d* only, and calibrations in *f* are for *f* only. The time calibration in *g* pertains to *e* and *g*. The voltage calibration in *g* pertains to *b*, *c*, *e* and *g*.

Methods: Wistar rats were deeply anaesthetized with sodium pentobarbital (50 mg per kg). Animals of various ages were used for developmental studies and 12-week-old rats were selected for regeneration studies. Their sciatic nerves were exposed and a discrete nerve crush was made in the mid-thigh with fine forceps for 20 s (see ref. 6). The incisions were closed and the rats allowed to recover for various post-crush times. At these times the nerves were removed, desheathed and placed in warmed oxygenated Ringer solution (124 mM NaCl, 3.0 mM KCl, 2.0 mM MgSO₄, 2.0 mM CaCl₂, 26 mM NaHCO₃ and 10 mM dextrose, bubbled with 95% O₂ and 5% CO₂). The nerves were then positioned in a modified sucrose gap chamber (*a*). The regenerated nerve segment to be studied was stimulated with bipolar Teflon-coated stainless steel stimulating electrodes. The compartment of the chamber containing the nerve segment to be studied was continuously superfused with warmed (37 °C) Ringer solution into which drugs could be added. The central compartment was continuously washed with isotonic sucrose, and the other outer compartment (left compartment in *a*) was washed with the normal electrolyte solution containing isotonic KCl. Osmotic balance was maintained in the isotonic KCl solution by reducing the NaCl concentration. Whole nerve recordings were obtained by recording with platinum black electrodes connected by agar/Ringer bridges to the outer compartments of the chamber. Leads from the electrodes were taken to inputs of a high impedance differential electrometer. Slow d.c. potential changes from the whole nerve bundle can be recorded across the sucrose gap¹⁶. The effectiveness of the gap chamber in recording d.c. potential changes was determined by introducing isotonic KCl into the test compartment; gap potentials gradually developed up to 60 mV after the addition of isotonic KCl. Intra-axonal recordings were obtained from impalements made within the test compartment (right compartment *a*) with glass microelectrodes filled with 2 M KCl (d.c. resistances 50–100 MΩ).



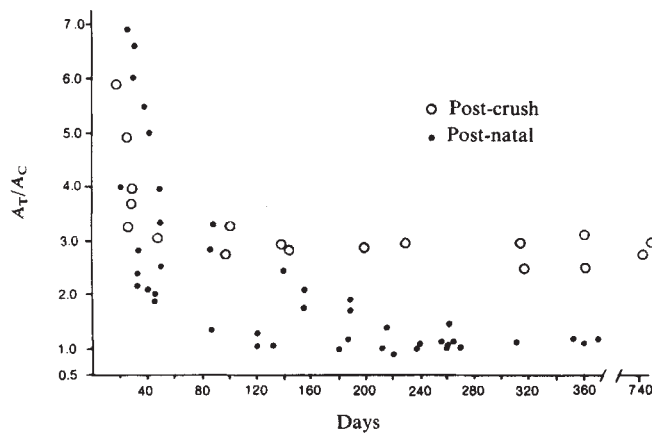


Fig. 2 Graph showing the amplitude of the compound action potential after 4-AP (1.0 mM; A_T) application divided by the response amplitude before 4-AP (A_C) application, as a function of age and duration of regeneration.

4-AP application no change in the gap potential was observed, indicating that 4-AP did not elicit a change in membrane potential.

Intra-axonal recordings were obtained from long-term (>1 yr post-crush) regenerating nerves before and after 4-AP application. None of the regenerated axons ($n = 18$) recorded from before 4-AP application gave rise to more than one action potential in response to a single stimulus. However, after 4-AP application, 83% of the regenerated axons ($n = 21$) gave rise to bursts of spikes. The intra-axonal recording in Fig. 1f was obtained from a regenerating axon (372 days post-crush) in the presence of 4-AP. Note the burst of spikes following a single whole nerve stimulus. The waveform of the initial spike is not noticeably broadened, but a delayed membrane depolarization develops on which the action potential burst arises. A hyperpolarization often follows the initial depolarization, as can be seen in Fig. 1f. The delayed depolarization and the bursts of spikes are the intra-axonal correlate of 4-AP-elicited late activity in the compound action potential recordings.

The emergence of late activity in the presence of 4-AP, which is due to delayed membrane depolarization and action potential burst activity following a single stimulus, is a feature shared by young myelinated axons and long-term regenerated myelinated axons; however, it is not a characteristic of normal mature myelinated axons. When the nerve segments proximal to the crush sites were studied, 4-AP had minimal influence on the waveform of the response (Fig. 1g). This further demonstrates that the influence of 4-AP is specific for the regenerating nerve segment.

Figure 2 shows the area of the compound action potential after 4-AP application (A_T) divided by the area of response before 4-AP application (A_C), plotted as a function of either postnatal or post-crush times. During the normal course of maturation, myelinated fibres initially have a large A_T/A_C ratio, but it approaches 1.0 as the animals mature⁵. By 200 days of age the ratio varies between approximately 1.0 and 1.5. In contrast, regenerating nerve fibres maintain an A_T/A_C ratio between 2.5 and 3.5 even at post-crush times of over 1 yr. It might be argued that normal nerve maturation begins at some time before birth. However, even if the results in Fig. 2 are compensated maximally for this, by adding 3 weeks (the gestation period in rat) to the postnatal time, the development of the regenerating fibres lags appreciably behind that of normal developing fibres. The A_T/A_C ratio remains high even at 2 yr after crush.

Voltage-clamp studies on mature mammalian myelinated axons indicate that voltage-dependent potassium currents (g_K)

are minimal at normal mature nodes⁷⁻⁹. But, upon demyelination a distinct late outward K current is evident, suggesting the presence of K channels at internodal axon regions^{7,9}. In agreement with these studies, external application of g_K blocking agents does not alter the waveform of the action potentials recorded from mature mammalian axons^{2,3}. However, during the course of maturation, young myelinated axons are exquisitely sensitive to 4-AP^{4,5} and they give rise to bursts of action potentials in response to single stimuli following blockage of g_K , thus indicating the importance of g_K in stabilizing nodal firing properties⁵.

One possibility to account for the diminution of the 4-AP effect during the course of normal maturation is that K channels are initially present at the node but redistribute or are otherwise lost during the course of maturation. If this is correct, then the persistent 4-AP effect in long-term regenerated axons indicates that the ability of the normal maturing axon to redistribute or delete K channels from the node is not shared by long-term regenerated myelinated axons. It is interesting, in this regard, that internode distances in regenerated fibres remain relatively short^{10,11} and that the nodal remodelling¹²⁻¹⁴ which occurs during normal development may not occur or may be attenuated during regeneration. Another possibility to explain the persistent 4-AP effect in long-term regenerated fibres is that the extra-axonal organization of the regenerated nerves allows for greater access to the peri-axonal (submyelinic) space for externally applied agents. Regenerated axons have thinner myelin and therefore fewer terminal myelin loops in the paranodal region^{10,11}. These morphological features are also present in immature normal myelinated nerve¹⁵. Rather than an inability of regenerated fibres to redistribute or delete nodal K channels, 4-AP may simply have better access to the peri-axonal space and therefore better access to K channels located at the internodal axolemma.

Conduction properties such as conduction velocity and the refractory period are not strikingly different in long-term regenerated axons compared with normal mature fibres^{4,10,11}. Furthermore, functional recovery is nearly complete. Yet, aminopyridine application leads to stimulus-evoked delayed membrane depolarization and action potential burst activity in long-term regenerated axons in a manner that is remarkably similar to the way aminopyridines affect immature myelinated fibres. Thus, the course of development of electrophysiological properties in regenerating fibres does not fully recapitulate that of normal myelinated fibres; rather, in regenerating fibres, functional maturation progresses to a stage similar to that of normal immature myelinated fibres. This result, taken together with known morphological features of regenerated axons, suggests that the functional organization of even long-term regenerated mammalian nerve fibres bears a closer resemblance to that of immature myelinated axons than to that of mature axons.

This work was supported in part by the Medical Research Service, Veterans Administration, and by grants from the National Multiple Sclerosis Society.

Received 14 March; accepted 31 May 1983.

1. Ramon Cajal. *Degeneration and Regeneration of the Nervous System* Vol. 1 (Oxford University Press, 1928).
2. Bostock, H., Sears, T. A. & Sherratt, R. M. *J. Physiol., Lond.* **313**, 301-315 (1981).
3. Kocsis, J. D. & Waxman, S. G. *Nature* **287**, 348-349 (1980).
4. Ritchie, J. M. *Proc. R. Soc. B215*, 289-297 (1982).
5. Kocsis, J. D., Ruiz, J. A. & Waxman, S. G. *J. Neurophysiol.* (in the press).
6. Kocsis, J. D., Waxman, S. G., Hildebrand, C. & Ruiz, J. A. *Proc. R. Soc. B217*, 77-87 (1982).
7. Brismar, T. *Acta physiol. scand.* **105**, 384-386 (1979).
8. Chiu, S. Y., Ritchie, J. M., Rogart, R. B. & Stagg, D. J. *J. Physiol., Lond.* **292**, 149-166 (1979).
9. Chiu, S. Y. & Ritchie, J. M. *J. Physiol., Lond.* **313**, 415-437 (1981).
10. Sanders, F. K. & Whitteridge, D. J. *J. Physiol., Lond.* **105**, 152-174 (1946).
11. Cragg, B. G. & Thomas, P. K. *J. Physiol., Lond.* **171**, 164-175 (1964).
12. Berthold, C. H. & Skoglund, S. *Acta. soc. med. upsal.* **73**, 127-144 (1968).
13. Friede, K. & Hildebrand, C. *J. Anat. Lond.* **134**, 517-531 (1982).
14. Waxman, S. G. & Foster, R. E. *Proc. R. Soc. B209*, 441-446 (1980).
15. Webster, H. de F. *J. Cell Biol.* **48**, 348-367 (1971).
16. Koketsu, K. & Nishi, S. *J. Physiol., Lond.* **196**, 293-310 (1968).

Selection and properties of a mouse L-cell transformant expressing human transferrin receptor

Roland Newman, Derrick Domingo,
Joseph Trotter & Ian Trowbridge

The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA

Transferrin receptors are expressed in large quantities on tissues with high requirements for iron such as maturing erythroid cells and placenta. In addition, they are found in abundance on proliferating cells from other normal tissues¹⁻⁵ as well as on a variety of tumours⁴⁻⁸. Recent genetic analysis has shown that structural genes for the transferrin receptor, probably transferrin itself and for p97, a melanoma-associated antigen that exhibits primary sequence homology with transferrin and that can bind ferric iron, each map in man to chromosome 3 (refs 9-12). On this basis it has been suggested that there may be a region on chromosome 3 containing genes involved in Fe transport and that rearrangements in this region of chromosome 3 may in some circumstances be associated with malignant transformation¹². Furthermore, it is unresolved whether all cell types express structurally identical transferrin receptors^{13,14}. To study these problems, and as an initial step towards cloning the transferrin receptor gene, we describe here the derivation of mouse L-cell transformants expressing the human transferrin receptor.

Mouse L-cell transformants expressing the human transferrin receptor on their cell surface were obtained as follows: mouse L (thymidine kinase-negative, tk⁻) cells were cotransformed with a recombinant plasmid containing the herpes simplex type 1 thymidine kinase (tk) gene¹⁵ and high molecular weight DNA isolated from the human T-cell line, CCRF-CEM¹⁶, using the calcium phosphate method¹⁷ as described in detail in Fig. 1 legend. Transformants containing the tk gene were selected by growth in Dulbecco's modified Eagle's medium (DMEM) containing 10% horse serum and hypoxanthine-aminopterin-thymidine (HAT) medium. In each trial, 2-4 × 10³ HAT-resistant colonies were pooled, stained with a mixture of mouse monoclonal antibodies against the human transferrin receptor followed by fluorescein isothiocyanate-conjugated goat-anti-mouse immunoglobulin and then sorted by flow cytometry. The brightest 5% of the viable fluorescent cell population was collected, grown up and resorted. From a total of 11 independent pools of HAT-resistant clones, two transformants were obtained that express the human transferrin receptor on their cell surfaces. The properties of one of these transformants, designated J4, is described below.

After three cycles of sorting a population was obtained that was more than 90% positive when stained with the B3/25 monoclonal antibody⁴ against the human transferrin receptor. Transformants were also positive when stained with anti-human transferrin receptor monoclonal antibodies 42/6 (ref. 18), OKT9 (refs 5, 19) and T56/14 (ref. 4). The level of expression of the murine transferrin receptor on J4 cells was comparable with that of L(tk⁻) cells and karyotypic analysis confirmed that the transformant was derived from the original L-cell population.

Figure 2 shows SDS-polyacrylamide gel electrophoresis (SDS-PAGE) of immunoprecipitates prepared from lysates of J4 cells labelled by lactoperoxidase-catalysed cell-surface iodination. The results confirm that the transformed cells express both murine and human transferrin receptors as shown by the precipitation of labelled species of molecular weight ~90,000 (90K)-100K in reducing conditions with both B3/25

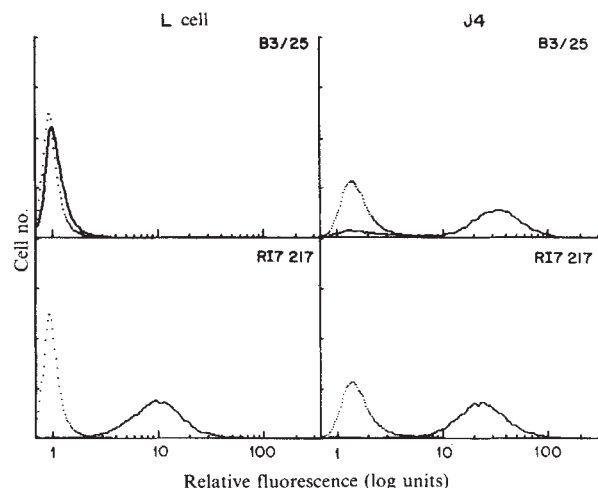


Fig. 1 Expression of human transferrin receptor on mouse L cells after transformation with human DNA followed by FACS. Solid line, cells stained with specific monoclonal antibody; dotted line, cells stained with second layer fluorescent antibody alone. Mouse L (tk⁻) cells²² (3×10^5 per 35 mm tissue culture plate) were incubated for 5 h with a calcium phosphate precipitate^{17,23} of 1 µg of a recombinant plasmid containing the herpes simplex type 1 tk gene¹⁵ and 7 µg of CCRF-CEM DNA which had been mechanically sheared by passing it 10 times through a 25-gauge needle. After incubation the cells were exposed to DMEM containing 15% glycerol for 1 min. Cells were then washed in DMEM and incubated for 48 h in DMEM containing 10% horse serum, then transferred to HAT medium. After 2-3 weeks, about 200 colonies were obtained from each original plate of 3×10^5 cells. Colonies ($\sim 4 \times 10^3$) from 20 plates were pooled, washed in HEPES-buffered Eagle's medium²⁴ containing 0.2% polyvinylpyrrolidone (PVP) and 2 mM sodium azide and incubated with the appropriate monoclonal antibody diluted in HEPES/2% fetal calf serum (FCS; 500 µl per 10^7 cells) for 30 min on ice. Cells were washed three times in HEPES/PVP/azide and incubated on ice for 30 min with highly fluoresceinated goat anti-mouse IgG (Antibodies Inc.) or goat anti-rat IgG (200 µl per 10^7 cells; Cappel Laboratories). After washing with HEPES/PVP/azide, cells were suspended (4×10^6 ml⁻¹) in HEPES/5% FCS, filtered through sterile nylon gauze and an equal volume of propidium iodide solution (10 µg ml⁻¹ in phosphate-buffered saline, PBS) added. Cells were then sorted using the Salk Institute cell sorter as described previously²⁴ except that the instrument was equipped with a Spectra-Physics 164-05 argon laser. The brightest 5% of the cells stained with the monoclonal antibody B3/25 were separated, grown in DMEM/10% horse serum/HAT and resorted. J4 cells were obtained by three sequential sorts of the original HAT-resistant transformants. Flow cytometric analysis was done on the Salk Institute flow microfluorimeter interfaced with a Digital Equipment Corporation LSI-11/23 computer. Analog signals from 20,000 cells per sample were collected and stored as list-mode data files on computer disk. Dead cells were gated out as previously described²⁵ and the remaining data plotted as relative fluorescence, determined by calibrating the instrument's logarithmic compression of integrated fluorescence with bead standards run at the same 488 nm laser output of 800 mW, high voltage, and logarithmic amplifier input sensitivity. The relative fluorescence value of 1 was assigned to the mean fluorescence of the L cells stained with fluorescent antibody alone. At the time of this analysis, 81% of the J4 cells were stained with B3/25 monoclonal antibody with a mean fluorescence 32 times higher than the mean background fluorescence. The instrument's modified LACEL²⁶ software converts any mean value to a relative value for the logarithmically compressed data and scales the histogram as shown.

monoclonal antibody and RI7 217, a rat monoclonal antibody against the murine transferrin receptor (ref. 20 and J. Lesley and R. Schulte, unpublished results). The murine and human transferrin receptors could be distinguished by SDS-PAGE on the basis of their apparent molecular weights, the murine receptor migrating significantly more slowly than the human receptor. Both species migrated as dimers of ~190K in non-reducing

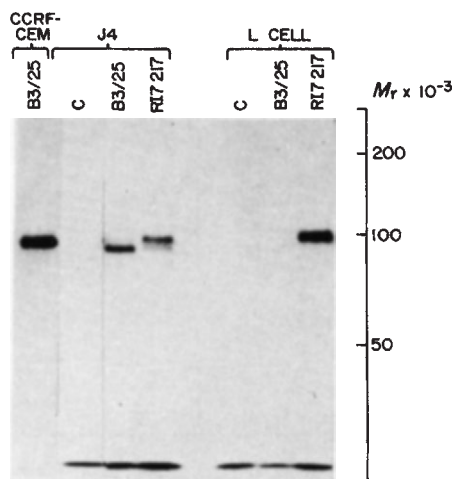


Fig. 2 SDS-PAGE of murine and human transferrin receptors recognized by the monoclonal antibodies RI7 217 and B3/25 respectively. L cells, J4 transformants and the lymphoid cell line CCRF-CEM were iodinated by the lactoperoxidase technique as previously described⁴. Cells were lysed in 10mM Tris pH 8.0, 0.15 M NaCl, 1% Nonidet P-40, 1mM EDTA, 1 mg ml⁻¹ bovine serum albumin (BSA) and immunoprecipitates prepared by incubating aliquots with the mouse anti-human transferrin receptor monoclonal antibody, B3/25, or the rat anti-murine transferrin receptor monoclonal antibody, RI7 217, at 4°C for 1 h. *Staphylococcus aureus* precoated with rabbit anti-mouse IgG or rabbit anti-rat IgG respectively was used to precipitate the immune complexes. Control slots using only rabbit anti-mouse IgG were also run (gel slot marked C). Dried gels were exposed to Kodak XAR-5 film for 24 h using Lightning Plus intensifying screens²⁷ (Dupont). M_r , molecular weight.

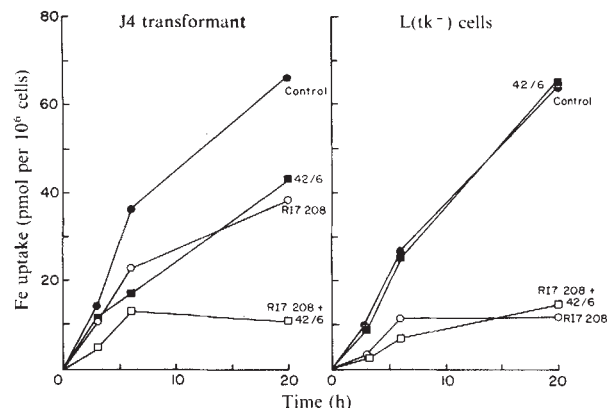


Fig. 3 Measurement of Fe uptake in the J4 transformant in the presence of monoclonal antibodies against human and murine transferrin receptors. ⁵⁹Fe-labelled human transferrin (specific activity 2×10^4 c.p.m. μg^{-1}) was prepared essentially as described by Bates and Schlabach²⁸ using the nitrilotriacetic acid method. L(tk⁻) and J4 cells were plated at 6×10^5 cells per 35 mm tissue culture dish in DMEM supplemented with 10% horse serum; the following day the cell monolayers were washed extensively with serum-free DMEM at 37°C and incubated for 30 min in 1 ml DMEM containing either 50 $\mu\text{g ml}^{-1}$ of purified monoclonal anti-human transferrin receptor (42/6), monoclonal anti-murine transferrin receptor (RI7 208), or both antibodies together. After preincubation with the antibodies, ⁵⁹Fe-human transferrin was added in 50 μl of DMEM to give a final concentration of 5 $\mu\text{g ml}^{-1}$. At 0 min and various times thereafter, duplicate dishes from each treatment group were washed three times with 2 ml of cold 0.15 M NaCl–0.015 M NaH₂PO₄–0.01 M Na phosphate buffer containing 0.1% BSA. The cells were then removed from the dish with versene and ⁵⁹Fe radioactivity was measured in a γ -counter.

conditions. Two further points of interest emerge from SDS-PAGE analysis. First, the human transferrin receptor precipitated by B3/25 from J4 cells is of lower apparent molecular weight than that from CCRF-CEM cells. Second, in addition to the major species precipitated by B3/25 and RI7 217 monoclonal antibodies, each antibody (most clearly seen in the case of RI7 217 in the experiment illustrated in Fig. 2) precipitated a minor band. These species were not observed when mouse and human transferrin receptors were isolated from L cells or CCRF-CEM cells respectively, thus they are probably a result of the formation of heterodimers between the murine and human polypeptides. Such heterodimers can be detected by passing cell extracts through a B3/25 monoclonal antibody-Sepharose column and analysing the transferrin receptor species that are either bound or pass through the column (I.T., unpublished results).

As monoclonal antibodies are available that specifically block the uptake of transferrin-bound Fe mediated by the human and mouse transferrin receptors respectively, it was possible to ask the question whether the human transferrin receptors expressed in the J4 transformants were functional. The data in Fig. 3 clearly establish that the human transferrin receptors expressed on J4 cells can transport transferrin-bound Fe across the cell membrane. Both the monoclonal antibody 42/6 that blocks transferrin binding to the receptor of human cells and monoclonal antibody RI7 208 that interferes with Fe uptake by the murine transferrin receptor partially inhibit Fe uptake in J4 cells. The two antibodies together, however, inhibit Fe uptake by J4 cells much more effectively. The degree of inhibition of Fe uptake in J4 cells obtained with a combination of 42/6 and RI7 208 antibodies is similar to that obtained if L(tk⁻) cells are treated with monoclonal antibody RI7 208 alone. As expected, monoclonal antibody 42/6 did not interfere with Fe uptake in L cells.

The results reported here demonstrate that it is possible by fluorescence-activated cell sorting (FACS) to obtain, from pools of L(tk⁻) cells cotransformed with herpes tk gene and high molecular weight human DNA, cells that express the human transferrin receptor. The human transferrin receptor is functional and expressed in amounts similar to the murine receptor itself. The J4 transformant described here was analysed over 3 months after the original transfection procedure was carried out, although it has now maintained stable expression of the human transferrin receptor for >6 months when grown in selective HAT medium. The second independent L-cell transformant expressing the human transferrin receptor is similar in this respect. Six independent stable clones have now been established that were >99% positive when stained with B3/25 monoclonal antibody. The expression of the human transferrin receptor in J4 cells appears to be regulated coordinately with cell growth. This is consistent with the genetic element(s) regulating transferrin receptor expression being closely associated with the structural gene. Further work will be required to determine the detailed mechanism of this genetic regulation. In addition to their utility for studies of the structure and genetic regulation of the transferrin receptor, J4 cells have proved to be a useful immunogen for the preparation of anti-receptor antibodies in C3H mice, the inbred strain of mice from which L cells were originally derived.

Recently, using similar methodology, Kavathas and Herzenberg²¹ have reported the selection of L-cell transformants expressing the human T-cell differentiation antigens, Leu 1 and Leu 2. Our results support their conclusion that cotransformation and cell sorting offers a viable general approach for the isolation of genes encoding cell-surface molecules that can be detected on the basis of their antigenicity.

This work was supported by NCI grant CA34787 and funds donated by the Armand Hammer Foundation, the Paul Stock

Foundation, the Helen K. and Arthur E. Johnson Foundation, BankAmerica Foundation, Mr and Mrs Lyndon C. Whitaker Charitable Foundation, Frances and Charles G. Haynsworth and Constance and Edward L. Grund. We acknowledge helpful discussions with Dr R. Hyman.

Received 21 March; accepted 23 June 1983.

- Larrick, J. W. & Creswell, P. J. *supramolec. Struct.* **11**, 576–586 (1979).
- Hamilton, T. A., Wada, H. G. & Sussman, H. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6406–6410 (1979).
- Galbraith, G. M. P., Galbraith, R. M. & Faulk, W. P. *Cell. Immun.* **49**, 215–222 (1980).
- Trowbridge, I. S. & Omary, M. B. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3039–3043 (1981).
- Sutherland, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4515–4519 (1981).
- Faulk, W. P., Hsi, B.-L. & Stevens, P. L. *Lancet* **II**, 390–392 (1980).
- Shindelman, J. E., Ortmeier, A. E. & Sussman, H. H. *Int. J. Cancer* **27**, 329–334 (1981).
- Gatter, K. C., Brown, G., Trowbridge, I. S., Woolston, R.-E. & Mason, D. Y. *J. clin. Path.* **36**, 539–545 (1983).
- Goodfellow, P. M. et al. *Somatic Cell Genet.* **8**, 197–206 (1982).
- Enns, C. A., Suomalainen, H. A., Gebhardt, J. E., Schroder, J. & Sussman, H. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3241–3245 (1982).
- Miller, Y., Seligman, P., Jones, C., Patterson, D. & Scoggin, C. *Clin. Res.* **30**, 41A (1982).
- Plowman, G. D. et al. *Nature* **303**, 70–72 (1983).
- Lebman, D. et al. *Blood* **59**, 671–678 (1982).
- Stein, B. S. & Sussman, H. H. *J. biol. Chem.* **258**, 2668–2673 (1983).
- Enquist, L. W., Vandewoude, G. F., Wagner, J. R., Smiley, J. R. & Summers, W. C. *Gene* **7**, 335–342 (1979).
- Adams, R. A. et al. *Expl. Cell Res.* **62**, 5–10 (1970).
- Graham, F. L. & Van Der Eb, A. J. *Virology* **52**, 456–467 (1973).
- Trowbridge, I. S. & Lopez, F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1175–1179 (1982).
- Reinherz, E. L., Kung, P. C., Goldstein, G., Levey, R. H. & Schlossman, S. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1588–1592 (1980).
- Trowbridge, I. S. & Domingo, D. L. *Cancer Surv.* **1**, 543–556 (1982).
- Kavathas, P. & Herzenberg, L. A. *Proc. natn. Acad. Sci. U.S.A.* **80**, 524–528 (1983).
- Kit, S., Dubbs, D., Piekarski, L. & Hsu, T. *Expl. Cell Res.* **31**, 297–312 (1963).
- Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725–731 (1978).
- Hyman, R., Trowbridge, I., Stallings, V. & Trotter, J. *Immunogenetics* **15**, 413–420 (1982).
- Trowbridge, I., Lesley, J., Schulte, R., Hyman, R. & Trotter, J. *Immunogenetics* **15**, 299–312 (1982).
- Salzman, G., Wilkins, S. & Whitfill, J. *Cytometry* **1**, 337–341 (1981).
- Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82**, 314–316 (1977).
- Bates, G. W. & Schlaback, M. R. *J. biol. Chem.* **218**, 3228–3232 (1973).

Na⁺/H⁺ exchange and cytoplasmic pH in the action of growth factors in human fibroblasts

W. H. Moolenaar, R. Y. Tsien*,
P. T. van der Saag & S. W. de Laat

Hubrecht Laboratory, International Embryological Institute,
Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

The mechanisms by which growth factors stimulate metabolism and cell proliferation are largely unknown. Recent evidence suggests that mitogens rapidly activate a Na⁺/H⁺ exchange mechanism in the plasma membrane of their target cells^{1–3}, implicating cytoplasmic pH (pH_i) as a potential 'messenger'. Indeed, growth stimulation of quiescent fibroblasts leads to intracellular alkalinization at ~1 h after mitogen addition, as measured by weak-acid distribution methods^{4,5}. We have used an internalized fluorescent pH_i indicator⁶ to examine the pH_i-regulating mechanisms in diploid human fibroblasts and to obtain the first continuous pH_i recordings of the response to growth factors. We report here that (1) pH_i in human fibroblasts is controlled by a membrane-bound Na⁺/H⁺ exchanger, which rapidly restores pH_i after an acute cytoplasmic acidification, and (2) epidermal growth factor (EGF) and serum factors induce a rapid and persistent elevation of pH_i by modifying the pH_i sensitivity of the Na⁺/H⁺ exchanger. We conclude that in addition to having a basic role in pH_i regulation, Na⁺/H⁺ exchange may function as a transmembrane signal transducer in the action of growth factors.

Early passage human foreskin fibroblasts (HF cells), grown to confluence on rectangular glass coverslips, were loaded with the pH-sensitive indicator 2',7'-bis(carboxyethyl)-5,6-carboxy-fluorescein (BCECF) by uptake of its membrane-permeable

ester⁶, as outlined in Fig. 1 legend. About 10⁵ cells of a BCECF-labelled monolayer were illuminated in a spectrofluorimeter and the pH_i-dependent emission intensity was continuously recorded. Calibration of cytoplasmic dye fluorescence as a function of pH_i was obtained from H⁺ equilibration methods using the K⁺/H⁺ ionophore nigericin as previously described^{6,7} and briefly shown in Fig. 1a.

The average pH_i of confluent, quiescent HF cells in HCO₃⁻-free HEPES-buffered medium (HBS, pH 7.40) is estimated at 7.05 ± 0.02 (n = 8). We first examined the pH_i-regulating mechanisms in HF cells by analysing the recovery of pH_i from a sudden acid load, as induced by an NH₄⁺-prepulse^{8,9} (Fig. 1b). Here we summarize our results, the details of which will be presented elsewhere (W.H.M., L.G.J.T., R.Y.T., P.T.vdS., S.W.dL., in preparation). Following an NH₄⁺-acid load, pH_i recovery follows an exponential time course and is 90% complete in 6–7 min (rate constant 0.38 min⁻¹ ± 0.04, n = 8). The following results indicate that pH_i recovery is mediated by Na⁺/H⁺ exchange: (1) pH_i recovery is reversibly blocked by amiloride, an inhibitor of Na⁺/H⁺ exchange^{10–14} (half-maximal effect at 0.1 mM, Fig. 1c); (2) pH_i recovery is accompanied by an increase in amiloride-sensitive ²²Na⁺ influx and H⁺ efflux (not shown); (3) the rate of pH_i recovery depends on the external Na⁺ concentration ([Na⁺]_o) (Fig. 1d); half-maximal rate occurs at 30–40 mM Na⁺, while saturation is observed at ~100 mM.

Figure 1 e–g illustrates that Na⁺/H⁺ exchange may be reversed by imposing an appropriate in > out Na⁺ gradient: exposing the cells to Na⁺-free medium induces a steep fall in pH_i to ~6.4, accompanied by a net uptake of external H⁺. Both events are amiloride sensitive, consistent with the fact that they represent the reversed mode of the Na⁺/H⁺ exchanger. When Na⁺ is added back, pH_i immediately recovers, accompanied by net H⁺ extrusion; again, amiloride is inhibitory. Finally, we recall that, if one reasonably assumes that in HCO₃⁻-free medium the cytoplasmic buffering power is independent of pH_i, the rate of H⁺ extrusion is proportional to d(pH_i)/dt (refs 9, 15). The exponential behaviour of pH_i recovery then implies that the Na⁺/H⁺ exchange rate increases linearly as pH_i is reduced.

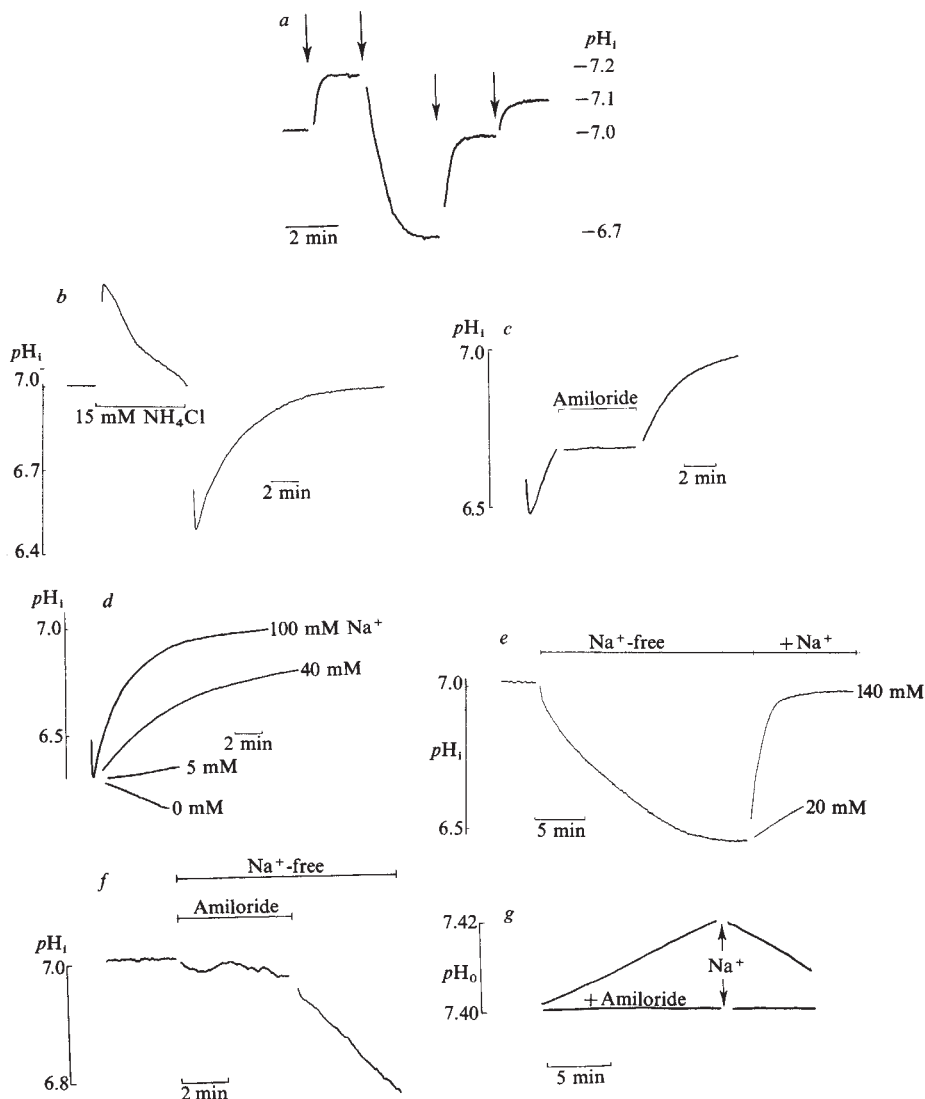
Next, we examined the response of quiescent HF cells to growth stimulation. Figure 2a shows that dialysed fetal calf serum (FCS, 10% v/v) raises pH_i by ~0.2 pH unit within 15–20 min. The FCS-induced pH_i shift varied between 0.17 and 0.25 pH unit in nine different experiments (average shift 0.21 ± 0.01). Because depleted FCS, which lacks mitogenic activity, has a negligible effect on pH_i, we conclude that the observed rise in pH_i is specifically induced by growth factors present in FCS. The polypeptide growth factor EGF (25 ng ml⁻¹) causes pH_i to increase by ~0.1 pH unit after 20 min (average shift 0.09 ± 0.01, n = 4), as shown in Fig. 2b. In the continuous presence of EGF or FCS, pH_i remains elevated for at least 1.5 h, whereas it slowly declines after mitogen washout; actually, pH_i recovery is not complete until 1–2 h after EGF or FCS withdrawal.

If the mitogen-induced pH_i rise is mediated by the Na⁺/H⁺ exchanger, it should be inhibited by amiloride. This is indeed observed (Fig. 2c). Interestingly, when FCS and amiloride are both washed out, pH_i rapidly increases and overshoots the original resting level by ~0.1 pH unit, as if amiloride had not been present during FCS application. This clearly shows that amiloride does not interfere with the mechanisms underlying the putative activation of Na⁺/H⁺ exchange by FCS. Similar effects, although less pronounced, were observed with EGF. Direct activation of Na⁺/H⁺ exchange by growth factors would also predict that the resultant rise in pH_i is converted to a fall in pH_i when the H⁺ driving force is reversed, as, for example, on removal of external Na⁺. This prediction is confirmed in Fig. 2d, where FCS is seen to increase the rate of cytoplasmic acidification in Na⁺-free medium. From these findings, together with earlier data on mitogen-induced Na⁺ influx in HF cells^{2,16}, we conclude that growth factors like EGF and serum factors

* Department of Physiology-Anatomy, University of California, Berkeley, California, 94720, USA.

Fig. 1 *a*, Calibration of the fluorescence intensity of intracellularly trapped 2',7'-bis(carboxyethyl)-5,6-carboxyfluorescein (BCECF) versus pH_i in confluent HF cells. *b*, Fluorimetric recording of the time course of pH_i recovery from an acid load induced by transient application of 15 mM NH_4Cl . Briefly, exposure of the cells to NH_4^+ results in a rapid intracellular alkalization due to NH_3 entry, while a small influx of NH_4^+ causes pH_i to decay slowly. When external NH_4^+ is removed, the accumulated NH_3 is forced to leave the cells as NH_4^+ , thereby loading the cells with an excess of intracellular protons (for full details see refs 8, 9). Media were changed by rapid perfusion. Temperature, 30–32 °C. External pH , 7.40. *c*, Effect of transient exposure to 1 mM amiloride on pH_i recovery. At 30 °C, amiloride (a gift from Merck, Sharp & Dohme) alone acidifies the cytoplasm at a rate of ~ 0.05 pH unit per 10 min. *d*, Time course of pH_i recovery at different external Na^+ concentrations. BDA $^+$ (bis(2-hydroxyethyl)dimethylammonium) was used as Na^+ substitute; choline gave similar results, but was cytotoxic in long-term experiments. *e*, Effect of 0 mM Na^+ (replaced by BDA $^+$) on pH_i and subsequent readdition of Na^+ . *f*, Same experiment as in *e* (first part), except for transient application of 0.5 mM amiloride. *g*, External pH changes induced by dilution of suspended cells ($\sim 5 \times 10^6$ cells ml^{-1}) into weakly buffered Na^+ -free medium (glucose-free; Na^+ replaced by choline). Arrow indicates addition of NaCl (final concentration, 20 mM). Amiloride was present at 0.5 mM. Full details in ref. 12.

Method for *a*: Quiescent HF monolayers, attached to a 3×1 cm glass coverslip, were loaded with BCECF by incubating them for ~ 1 h at 37 °C in HEPES-buffered medium (HBS, pH 7.10) containing 10 μM of the acetoxymethyl (AM)-ester of BCECF (ref. 6). Impermeant dye is then synthesized *in situ* by the action of cytoplasmic esterases. In HF monolayers, BCECF is almost exclusively located in the cytosol, as evidenced by digitonin-induced cell lysis and fluorescence microscopy. HBS composition in mM: 140 NaCl, 5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 10 HEPES, 10 glucose. After washing, the BCECF-loaded monolayers were inserted in a thermostatically controlled cell holder in a Perkin-Elmer 3000 fluorescence spectrometer (excitation wavelength 500 nm; emission wavelength 530 nm; slits 5×10 nm). Cells were incubated in 3 ml HBS (pH 7.40) at 30–32 °C. At the indicated times (arrows) the test medium was rapidly and completely replaced by a 150 mM KCl-HEPES buffer (prewarmed to 30 °C) of varying pH (as indicated) containing 10 μM of the K^+/H^+ ionophore nigericin. This results in an approximate equilibration of external and internal pH (see refs 6, 7), allowing calibration of fluorescence intensity against pH. The linear relationship between fluorescence intensity and pH over the pH range 6.40–7.60 (pK_a of BCECF ~ 7.0 ; ref. 6) was verified. Background fluorescence intensity was negligible. BCECF/AM was synthesized by one of us (R.Y.T.) as described recently⁶.



rapidly activate the previously 'quiescent' Na^+/H^+ exchanger, and thereby raise pH_i .

Of particular interest is the effect of insulin. This hormone is known to act synergistically with growth factors like EGF in stimulating DNA synthesis. Figure 2*e* illustrates that 5 μg ml^{-1} insulin alone fails to induce a detectable pH_i shift, but that it markedly potentiates the effect of EGF on pH_i (Fig. 2*f,g*). The combined action of insulin and EGF induces an alkaline pH_i shift of ~ 0.15 pH unit within 20 min, independent of the sequence in which the hormones are added. Table 1 summarizes the effects of various stimulants on pH_i .

By which mechanism(s) do growth factors activate Na^+/H^+ exchange? A clue is provided by kinetic analysis of pH_i regulation in both growth-stimulated and quiescent cells. As shown in Fig. 3*a*, the rate of exponential pH_i recovery after a NH_4^+ -acid load remains unaltered after growth stimulation (rate constant 0.35 min^{-1}), but the value at which pH_i stabilizes is ~ 0.2 pH unit more alkaline. This is expressed more quantitatively in Fig. 3*b*, where Na^+/H^+ exchange activity is measured as

$d(pH_i)/dt$ and plotted against pH_i . It is seen that growth stimulation simply induces a shift of the function $d(pH_i)/dt$ versus pH_i along the pH_i axis towards more alkaline values. In other words, growth factors make the Na^+/H^+ exchanger more sensitive to internal H^+ .

In experiments similar to those depicted in Fig. 1*d*, we have not detected FCS-dependent changes in the relationship between pH_i recovery rate and $[Na^+]_o$. Recent evidence¹⁷ suggests that the pH_i sensitivity of Na^+/H^+ exchange reflects allosteric behaviour with respect to internal H^+ . Thus, the simplest molecular model to explain our findings is that growth factors induce a conformational change at the allosteric site of the Na^+/H^+ exchanger, resulting in an increased apparent affinity for cytoplasmic H^+ . It is tempting to speculate that such a conformational change might be brought about by a mitogen-induced phosphorylation reaction^{18,19}.

Several authors have claimed a regulatory role for cytoplasmic Ca^{2+} in stimulating Na^+/H^+ exchange^{16,20,21}. This interpretation is mainly based on the finding that the Ca^{2+} ionophore

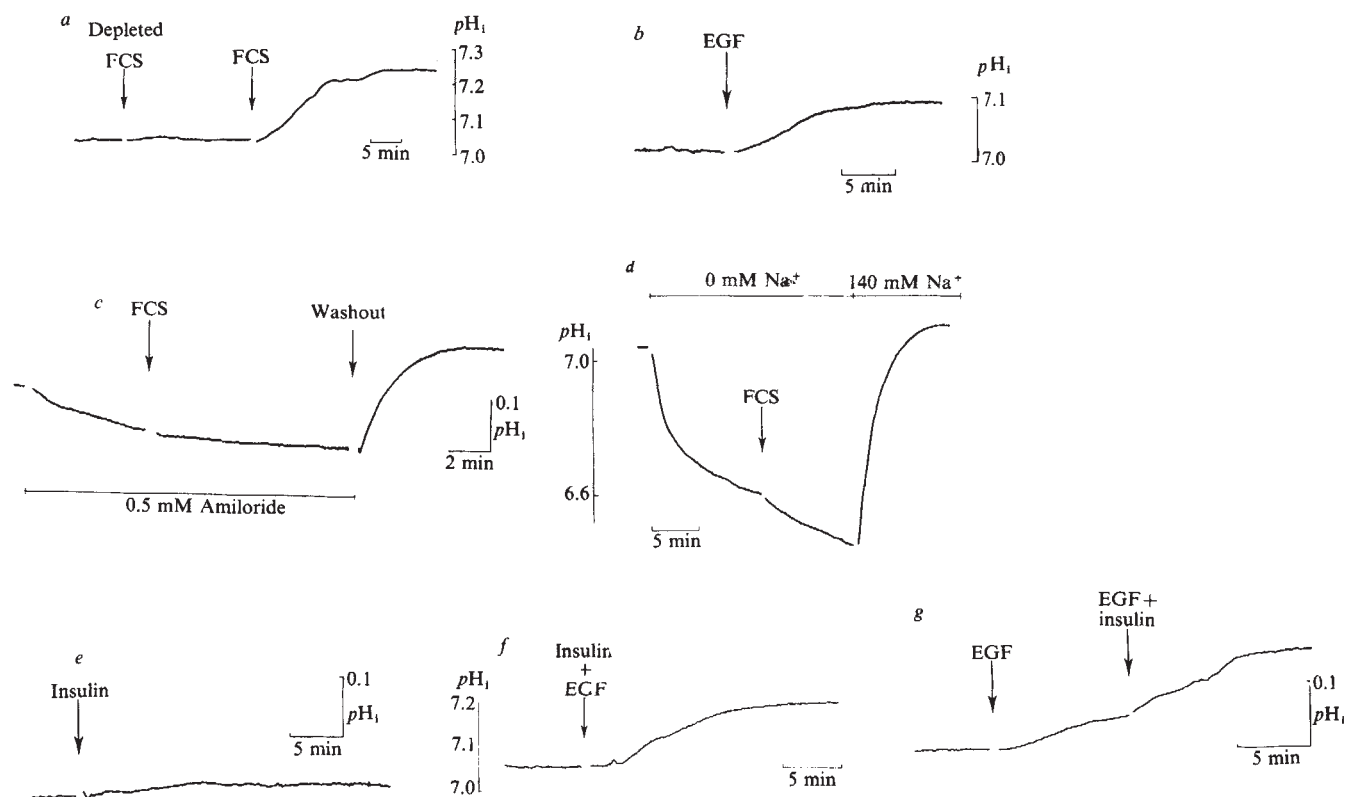


Fig. 2 Time course of pH_i changes induced by various mitogens. Quiescent cells were loaded with BCECF as in Fig. 1. Temperature, 30–32 °C. External pH , 7.40. *a*, Effect of growth factor-depleted FCS and fresh FCS (10% v/v), respectively. Depleted FCS was prepared as described in ref. 1. FCS was dialysed against HBS to remove weak acids and bases. *b*, Effect of 25 ng ml⁻¹ EGF (Collaborative Research, Inc.). *c*, Temporary inhibition of FCS-induced pH_i shift during transient application of 0.5 mM amiloride. *d*, Effect of 10% dialysed FCS at 0 mM and 140 mM Na⁺ respectively. Na⁺ was replaced by BDA⁺. *e*, Lack of effect of insulin (5 µg ml⁻¹) on pH_i . *f*, Potentiating effect of insulin on EGF-induced pH_i shift. Insulin (5 µg ml⁻¹) was added 20 min before addition of EGF (25 ng ml⁻¹). *g*, Same experiment as in *f*, except for reversed sequence of hormone addition.

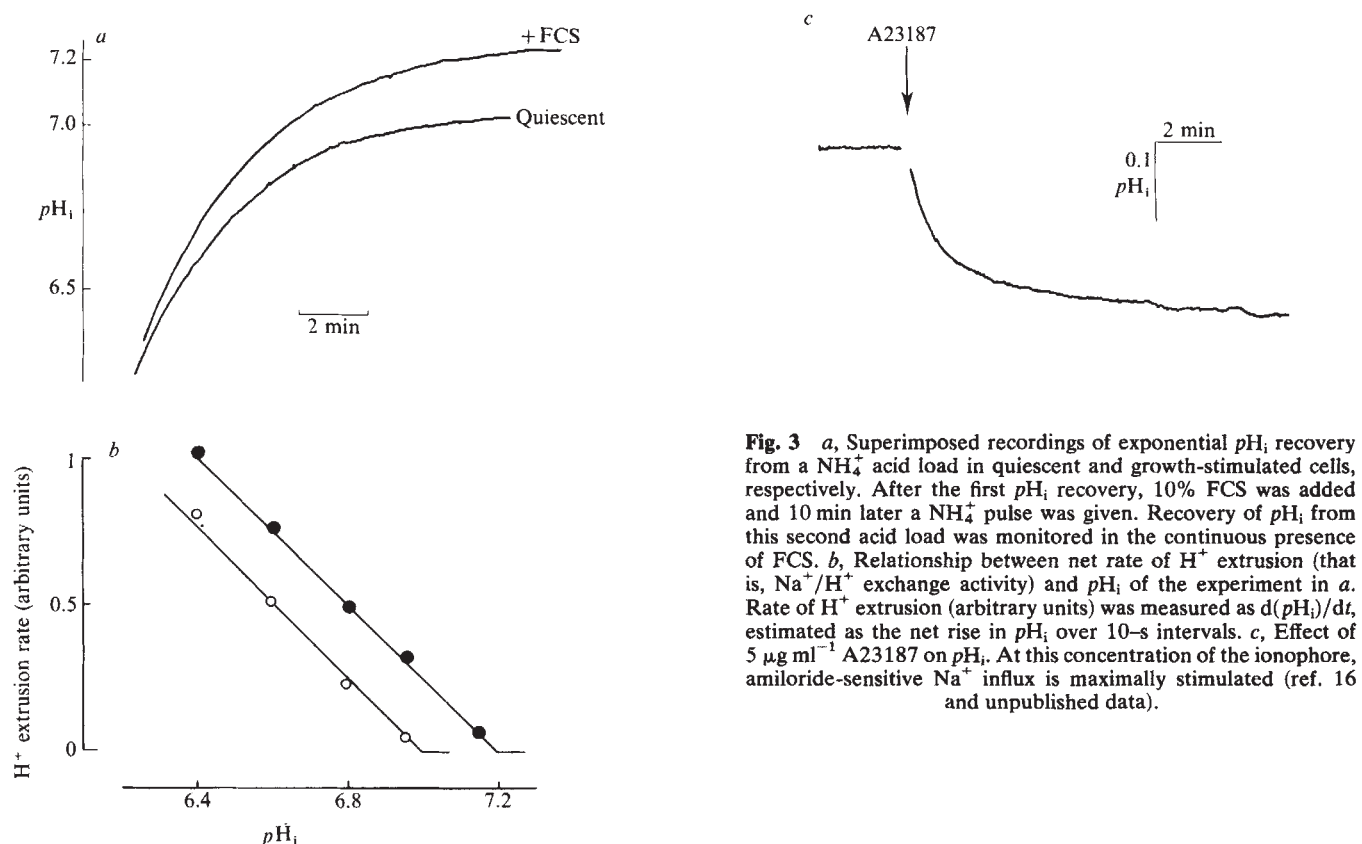


Fig. 3 *a*, Superimposed recordings of exponential pH_i recovery from a NH_4^+ acid load in quiescent and growth-stimulated cells, respectively. After the first pH_i recovery, 10% FCS was added and 10 min later a NH_4^+ pulse was given. Recovery of pH_i from this second acid load was monitored in the continuous presence of FCS. *b*, Relationship between net rate of H^+ extrusion (that is, Na^+/H^+ exchange activity) and pH_i of the experiment in *a*. Rate of H^+ extrusion (arbitrary units) was measured as $d(pH_i)/dt$, estimated as the net rise in pH_i over 10-s intervals. *c*, Effect of 5 µg ml⁻¹ A23187 on pH_i . At this concentration of the ionophore, amiloride-sensitive Na^+ influx is maximally stimulated (ref. 16 and unpublished data).

Table 1 Effects of various stimulants on pH_i

Conditions	pH_i	ΔpH_i (relative to control)
Quiescent control (HBS, pH 7.40, 30 °C)	7.05 $\pm$ 0.02 ($n = 8$)	
10% FCS	7.22–7.30	0.21 $\pm$ 0.01 ($n = 9$)
25 ng ml ⁻¹ EGF	7.14	0.09 $\pm$ 0.02 (4)
5 μ g ml ⁻¹ insulin		<0.01 (10)
EGF + insulin	7.20	0.15 $\pm$ 0.01 (6)
A23187 (5 μ g ml ⁻¹)	6.90	-0.14 $\pm$ 0.02 (5)

Quiescent HF cultures were loaded with BCECF and assayed for pH_i and shifts in pH_i (ΔpH_i) as detailed in the text and in Fig. 1 legend. The number of experiments is given in parentheses. The incubation period varied from 5 min (A23187) to 20–30 min (growth factors).

A23187 stimulates amiloride-sensitive Na^+ uptake. We have confirmed this result for HF cells (not shown). Figure 3c shows, however, that A23187 (5 μ g ml⁻¹) induces an abrupt and persistent fall in pH_i , perhaps due to enhanced mitochondrial Ca^{2+}/H^+ exchange²². Whatever the precise mechanism of this cytoplasmic acidification, we feel that the apparent activation of Na^+/H^+ exchange by A23187 is actually the result of a prior fall in pH_i and that there is no need to postulate a direct role for internal Ca^{2+} in the control of Na^+/H^+ exchange activity.

What is the physiological significance of a mitogen-induced rise in pH_i ? A small alkaline pH_i shift may accelerate such diverse pH-sensitive processes as glycolysis^{23,24}, protein synthesis^{25,26} and cytoskeletal reorganization^{27,28}. Our preliminary data indicate that the stimulation of glycolysis by growth factors can indeed be mimicked by artificially raising pH_i using the Na^+/H^+ ionophore monensin (W.H.M. and L. Joosen, unpublished results). Thus, the physiological implications of a rise in pH_i , mediated by Na^+/H^+ exchange, are potentially very significant. As such, Na^+/H^+ exchange may act as a signal transducer in the action of growth factors.

We thank Lilian Joosen and Leon Tertoolen for technical assistance. This work was supported by grants from the Netherlands Cancer Foundation (Koninklijke Wilhelmina Fonds).

Note added in proof: The FCS-induced pH_i shift appears to be largely due to platelet-derived growth factor (PDGF): 20 ng ml⁻¹ PDGF (provided by C.-H. Heldin) rapidly raises pH_i by ~ 0.15 unit, while the effects of PDGF and FCS are not additive (W.H.M., unpublished data).

Received 13 May; accepted 22 June 1983.

1. Moolenaar, W. H., Mummery, C. L., van der Saag, P. T. & de Laat, S. W. *Cell* **23**, 789–798 (1981).
2. Moolenaar, W. H., Yarden, Y., de Laat, S. W. & Schlessinger, J. *J. biol. Chem.* **257**, 8502–8506 (1982).
3. Pouyssegur, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3935–3939 (1982).
4. Schuldinger, S. & Rozenfurt, E. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7778–7782 (1982).
5. Moolenaar, W. H., de Laat, S. W., Mummery, C. L. & van der Saag, P. T. in *Ions, Cell Proliferation and Cancer* (eds McKeehan, W. L. & Boynton, A. L.) 152–162 (Academic, New York, 1982).
6. Rink, T. J., Tsien, R. Y. & Pozzan, T. *J. Cell Biol.* **95**, 189–196 (1982).
7. Thomas, J. A., Buschbaum, R. N., Zimniak, A. & Racker, E. *Biochemistry* **18**, 2210–2218 (1979).
8. Boron, W. F. & de Weer, P. *J. gen. Physiol.* **67**, 91–112 (1976).
9. Roos, A. & Boron, W. F. *Physiol. Rev.* **61**, 296–434 (1981).
10. Johnson, J. D., Epel, D. & Paul, M. *Nature* **262**, 661–664 (1976).
11. Aickin, C. C. & Thomas, R. C. *J. Physiol., Lond.* **273**, 295–316 (1977).
12. Moolenaar, W. H., Boonstra, J., van der Saag, P. T. & de Laat, S. W. *J. biol. Chem.* **256**, 12883–12887 (1981).
13. Kinsella, J. L. & Aronson, P. S. *Am. J. Physiol.* **241**, F374–F379 (1981).
14. Vigne, P., Frelin, C. & Lazdunski, M. *J. biol. Chem.* **257**, 9394–9400 (1982).
15. Thomas, R. C. *J. Physiol., Lond.* **273**, 317–338 (1977).
16. Villereal, M. L. *J. cell. Physiol.* **107**, 359–369 (1981).
17. Aronson, P. S., Nee, J. & Suhm, M. *Nature* **299**, 161–163 (1982).
18. Carpenter, G., King, L. & Cohen, S. *J. biol. Chem.* **254**, 4884–4891 (1979).
19. Ek, B. & Heldin, C.-H. *J. biol. Chem.* **257**, 10486–10492 (1982).
20. Taub, M. & Saier, M. H. *J. biol. Chem.* **254**, 11440–11444 (1979).
21. Benos, D. J. *Am. J. Physiol.* **242**, C131–C145 (1982).
22. Meech, R. W. & Thomas, R. C. *J. Physiol., Lond.* **265**, 867–879 (1977).
23. Trivedi, B. & Danforth, W. H. *J. biol. Chem.* **241**, 4110–4112 (1966).
24. Fidelman, M. L., Seeholzer, S. H., Walsh, K. B. & Moore, R. D. *Am. J. Physiol.* **242**, C87–C93 (1982).
25. Grainger, J. L., Winkler, M. M., Shen, S. S. & Steinhardt, R. A. *Dev. Biol.* **68**, 396–406 (1979).
26. Winkler, M. M., Steinhardt, R. A., Grainger, J. L. & Minning, L. *Nature* **287**, 558–560 (1980).
27. Regula, C. S., Pfeiffer, J. R. & Berlin, R. D. *J. Cell Biol.* **89**, 45–53 (1981).
28. Condeelis, J. & Vahey, M. *J. Cell Biol.* **94**, 466–471 (1982).

Fibroblast immortality is a prerequisite for transformation by EJ c-Ha-ras oncogene

Robert F. Newbold & Robert W. Overell

Chemical Carcinogenesis Section, Institute of Cancer Research, Pollards Wood Research Station, Nightingales Lane, Chalfont St Giles, Buckinghamshire HP8 4SP, UK

The established mouse cell line NIH 3T3 has been used with considerable success over the past three years as the basis of an *in vitro* transformation assay for demonstrating the presence of transfectable transforming genes in the DNA of certain human and rodent tumour cells (for review see ref. 1). In the case of the human bladder carcinoma cell lines EJ and T24, this approach has led^{2–4} to the molecular cloning of a transforming gene which is closely related to the rat-derived Harvey sarcoma virus oncogene, v-Ha-ras^{5,6}. A single point mutation, which distinguishes these genes from their normal human homologue (c-Ha-ras1), is thought to be solely responsible for their transforming potential^{7–10}. However, carcinogenesis in both humans and laboratory rodents is a multi-stage process (reviewed in ref. 11) of which the NIH 3T3 cell, already partly transformed, may represent only the penultimate stage. We therefore chose to examine the transforming effects of the EJ oncogene in a hamster fibroblast system originally developed in our laboratory to study stages in carcinogen-induced malignant transformation of normal diploid cells¹². We show here that EJ c-Ha-ras-1 lacks complete transforming activity when transfected into normal fibroblasts which have a limited lifespan, but can fully transform fibroblasts that have been newly 'immortalized' by carcinogens.

Primary diploid hamster dermal fibroblasts will reproducibly undergo approximately 20 population doublings (pd) in culture before losing their proliferative potential. During the early part of their culture lifespan they have a doubling time of 15–18 h in medium supplemented with fetal calf serum (FCS), and possess a high cloning efficiency (>40%) on tissue culture plastic. Unlike mouse fibroblasts, the cells are stable in culture, in that progeny capable of infinite multiplication do not arise spontaneously at a detectable frequency. Rare immortal variants can, however, be induced during the active growth phase (up to 15 pd) by treatment with carcinogens; these variants are usually observed as colonies (10^{-6} – 10^{-7} per treated cell) of rapidly dividing cells against a senescent background, at the end of the normal lifespan. We have proposed that this event may be an important early step in carcinogen-induced malignant transformation¹².

Newly immortalized hamster fibroblasts are invariably non-malignant in our inbred hamsters and, other than their unlimited lifespan, display normal growth characteristics in culture. For example, they depend on a high level (>10%) of FCS in medium for good growth, they reach a low saturation density in monolayer culture, and are unable to form colonies in soft agar (Aga⁻). The majority of lines will, however, acquire malignant properties on further subculture and such a transition is always marked by the appearance of anchorage-independent (Aga⁺) cells in the population. This phase in malignant transformation, which we term 'progression', occurs after an interval of subculture that varies between cell lines but which is a reproducible characteristic of each individual line¹².

To allow large numbers of cells to be screened more conveniently in studies of progression, we have recently developed a focus assay in which Aga⁺ clones are selected from confluent mass cultures in low (3%) serum. By comparing in parallel the agar and focus assays, plus the anchorage phenotypes of resulting transformants, we have shown that efficient recovery of Aga⁺ cells as foci is possible using the latter technique. It was

therefore chosen as the basis of the present study with the EJ oncogene.

In the first series of experiments, the standard calcium phosphate co-precipitation method¹³ was used to introduce EJ c-Ha-ras1 cloned^{3,7} in pBR322 (pEJ) into normal early-passage Syrian hamster dermal (SHD) cells and four newly immortalized cell lines. The latter were selected as being representative of the three distinct classes of immortal, premalignant cells so far observed in our laboratory (see Table 1 legend). In a focus assay, the following frequencies of spontaneous Aga⁺ foci were observed in control groups which had received carrier DNA alone: SHD, 4DH2 and 4XH2, $<10^{-7}$; 4XH1, $1.0 \pm 0.7 \times 10^{-6}$; 3M1, $3.4 \pm 1.5 \times 10^{-6}$. These were not significantly different from frequencies obtained in previous progression studies not involving DNA transfer. In all four cell lines (Table 1) pEJ proved extremely potent at inducing focus formation (>1 transformation event per ng plasmid DNA). Moreover, the susceptibility of each line to pEJ-induced transformation appeared to be unrelated to time in culture following immortalization, or to its tendency to progress spontaneously. Transformed foci were well defined against the flat density-inhibited background monolayer but were widely variable in size (Fig. 1a-d). Nevertheless, all foci possessed a characteristic morphology, consisting of rapidly dividing, basophilic, fusiform cells with little evidence of density inhibition or parallel orientation; this was distinct from spontaneous foci (for example 4XH1) or those induced in stable lines (for example, 4DH2) by an additional carcinogen treatment. Representative examples of the smallest

(F1) and largest (F2) foci scored were isolated for confirmation of anchorage independence. All cells so derived produced rapidly growing colonies in 0.33% agar, with high plating efficiency. In our experience, such overt anchorage-independent growth has always proved to be a reliable marker of malignancy¹². Tumorigenicity studies with pEJ transformants are now being done.

We then attempted to transform the stable line 4DH2 (Table 1) with the EJ c-Ha-ras1 gene using total genomic EJ DNA (gEJ) as donor. Six foci (Table 1) were recovered from 5.5×10^7 cells assayed after gEJ transfection, whereas none was obtained from 6.4×10^7 cells which had been exposed to normal human and hamster fibroblast DNA. We have since found no evidence of spontaneous progression in this line in more than 2×10^8 cells tested. The morphology of gEJ foci was strikingly similar to that of foci induced by pEJ, although the former were rather slower growing and the cells plated with a lower efficiency in agar.

Preliminary semiquantitative Southern analysis of complete BamHI digests of DNA from gEJ-induced 4DH2 transformants, using pEJ as a probe, revealed a single strongly hybridizing 6.6-kilobase (kb) band not present in untreated 4DH2 cells, confirming that EJ c-Ha-ras1 was the gene responsible for transformation. An additional (pBR322) band was seen in the case of pEJ transformants. Based on the amount of DNA applied to the gel and the band intensity compared with a series of standards, we calculated (two separate determinations) mean EJ c-Ha-ras1 copy number values of 0.8 and 16.0 per cell for

Table 1 Transforming effects of EJ c-Ha-ras oncogene following transfection into tertiary cultures of SHD cells and carcinogen-induced immortal variants

Recipient cell	Passage no. following immortalization	No. of foci induced by total genomic EJ (gEJ) DNA/no. of dishes	% Soft agar plating efficiency of cells from gEJ foci ($\pm$ s.d.)	No. of foci induced by pEJ plasmid DNA/no. of dishes	% Soft agar plating efficiency of cells from pEJ foci ($\pm$ s.d.)
4DH2	4	6/55	F1 12.3 ($\pm$ 2.1) F2 10.7 ($\pm$ 1.6)	5,440/10	F1 41.2 ($\pm$ 2.3) F2 62.7 ($\pm$ 2.9)
4XH1	3	ND	—	3,372/10	F1 52.0 ($\pm$ 1.7) F2 87.7 ($\pm$ 3.4)
4XH2	5	ND	—	4,132/10	F1 34.3 ($\pm$ 6.0) F2 73.4 ($\pm$ 4.2)
3M1	10	ND	—	3,420/10	F1 19.6 ($\pm$ 1.7) F2 81.9 ($\pm$ 4.1)
SHD	Limited proliferative potential	ND	—	0/25	—

Primary cultures of dermal fibroblasts (SHD cells) were prepared from newborn Syrian hamster skins by collagenase disaggregation of isolated dermis; separation of the epidermis from the dermis was facilitated by floating skins overnight in 0.01% pronase (Sigma) at 4°C. Immortal SHD cell lines were derived by exposure of replicate cultures of SHD cells (passage 3) to carcinogens as described previously¹². Cells from primary immortal foci were removed from dishes (9 cm; Sterilin) by trypsinization, re-spread into a single dish and, when just confluent, subcultured into three dishes to provide sufficient cells for cryopreservation. By this method, we obtained an extensive series of newly immortalized, serum- and anchorage-dependent lines with widely variable propensities for spontaneous progression to anchorage-independent growth and malignancy¹². The four cell lines used here were chosen to represent the three broad classes of premalignant line observed to date. Thus, lines 4DH2 and 4XH2 (induced by 20 μ g ml⁻¹ dimethyl sulphate and X rays (300 rad), respectively) consist of highly stable, serum-dependent cells which have shown no evidence of spontaneous progression; 4XH1 (300 rad X rays) typifies those lines which generate rare focus-forming Aga⁺ variants early on in their history, thereafter attaining tumorigenicity as these variants rapidly dominate the line due to selective advantage; 3M1 (30 μ g ml⁻¹ *N*-methyl-*N*-nitrosourea) progresses through two distinct steps to anchorage independence, that is, cultures are initially taken over by an intermediate, small focus-forming variant before becoming frankly Aga⁺ as a result of a subsequent rare event. Used here at passage 10 after immortalization, the line had already become dominated by cells which grew to a higher saturation density (Fig. 1d) and which were able to form microcolonies (5–10 cells) in 0.33% agar. Lines 4DH2, 4XH2 and 4XH1 were used as early as technically possible after immortalization. Unless stated otherwise, cell culture was done in Dulbecco's modified Eagle's minimal essential medium (DMEM) supplemented with 15% FCS (Gibco). Routine mycoplasma tests carried out on all cell lines by Hoechst 33258 staining, invariably proved negative. High molecular weight DNA was isolated from EJ bladder carcinoma cells, SHD cells (passage 3) and adult human dermal fibroblasts (passage 7) essentially by the method of Smith *et al.*²¹. This method routinely gave DNA with $A_{260/280}$ of 1.8–1.9 and molecular weight $>30 \times 10^6$ as determined by mobility on 0.4% agarose gels alongside undigested λ DNA (NEB). Plasmid pEJ (pBR322 with a 6.6-kb BamHI insert³ including the complete functional human EJ bladder carcinoma c-Ha-ras1 oncogene) was obtained from Professor R. A. Weinberg (MIT) and propagated in *Escherichia coli* HB101. Milligram quantities of this plasmid were prepared by a standard cleared-lysate procedure³. Complete BamHI digestion of pEJ generated the expected two fragments (4.4 and 6.6 kb) as assayed on a 0.7% agarose gel with HindIII digested λ DNA fragments as markers. For transfection experiments, ampoules of cells were thawed at the passage number shown and cells inoculated into 175 cm² culture flasks (Nunc). After a single subculture, 3.5×10^6 cells of each type were seeded into replicate 175 cm² flasks containing 40 ml of medium. After 24 h ($\sim 5 \times 10^6$ cells per flask) 10 ml of medium were removed from each culture and DNA added in the form of a calcium phosphate co-precipitate, essentially as described by Wigler *et al.*¹³. Each flask received 3 ml of precipitate containing 60 μ g DNA (SHD/human fibroblast carrier DNA or total genomic EJ DNA). Where appropriate, 1 μ g ml⁻¹ pEJ was included with the carrier DNA. Cultures were incubated at 37°C for 18 h in the presence of co-precipitates, after which the medium was renewed; 24 h later the cells were removed from flasks with trypsin/EDTA and re-plated in 9 cm dishes at a density of 10^6 cells per dish. After 24 h the medium was again replaced, this time with DMEM containing 3% FCS; this medium was renewed every 3 days for 2–3 weeks. As reported by Smith *et al.*²¹, reduction of the serum concentration to 3% resulted in greater contrast between transformed foci and the normal cell background and, in our hands, also prevented sloughing of the monolayer often found with some cell types after prolonged incubation. At the end of the incubation period, most of the cultures were fixed in methanol and stained with methylene blue. Transformed foci were scored using a low-power dissecting microscope. Two examples of foci (denoted F1 and F2) representative of extremes of sizes scored were isolated from the remaining plates using stainless steel cloning cylinders and the cells grown in mass culture (DMEM+3% FCS). As soon as enough cells were available ($\sim 10^7$) each line was assayed for anchorage-independent growth by plating 10^3 or 10^5 exponentially growing cells in 0.33% Noble agar (Difco) into 5 cm dishes as described previously¹². The remaining cells were preserved in liquid N₂ for subsequent Southern analysis. ND, not done.

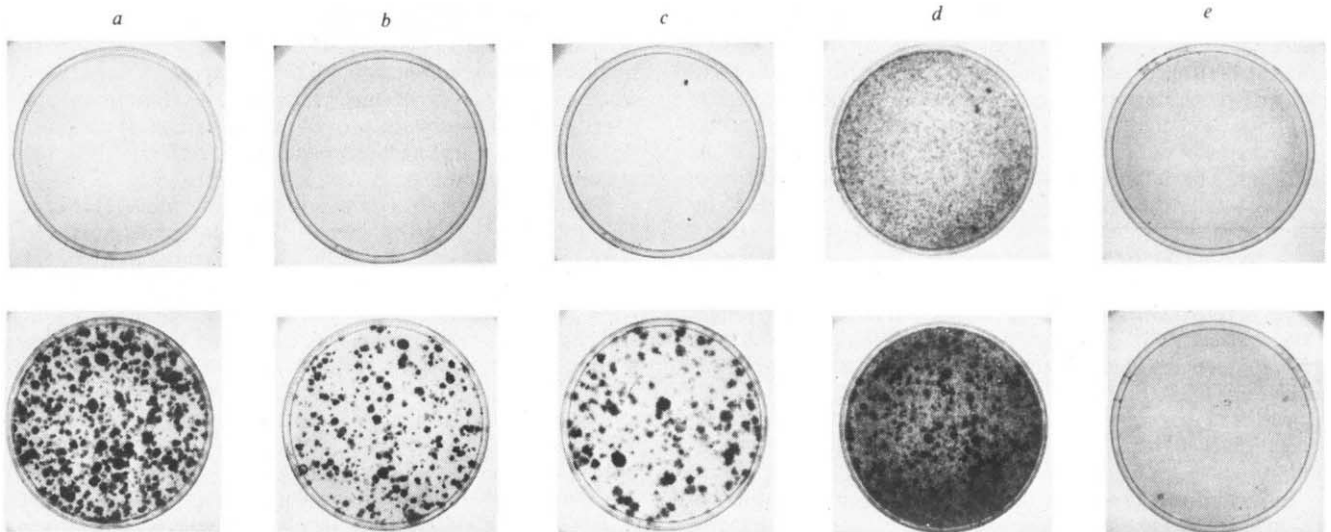


Fig. 1 Examples of culture dishes from experiments comparing the transforming effects of the EJ c-Ha-ras1 oncogene in tertiary SHD cells and four recently immortalized SHD cell lines: *a*, 4DH2; *b*, 4XH2; *c*, 4XH1; *d*, 3MI; *e*, SHD. Mass cultures of each type were transfected with pEJ and assayed for focus formation as described in Table 1 legend, except that the SHD cells shown here are those that had been exposed to a higher concentration of pEJ (15 μ g per flask). The upper dish in each case is the control (carrier DNA only). Cultures were fixed in methanol and stained with methylene blue 15 days after re-plating. $\times 0.32$.

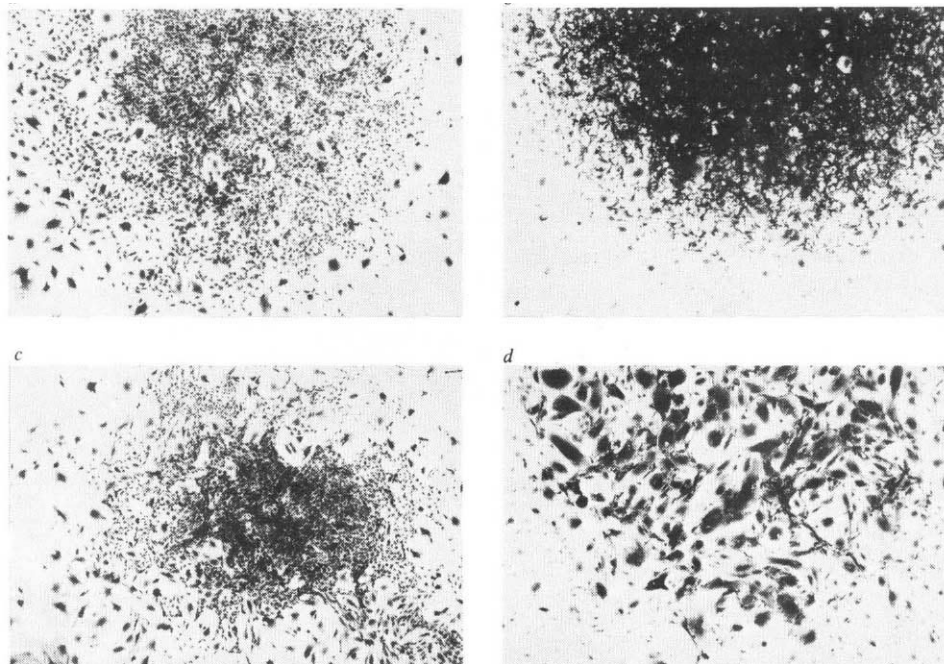


Fig. 2 Photomicrographs showing typical examples of Mtx^r colonies produced in pHG/pEJ co-transfer experiments: *a*, normal flat 4DH2 colony, pHG transfer only; *b*, transformed 4DH2 colony, pHG/pEJ co-transfer; *c*, normal Mtx^r SHD colony, pHG only; *d*, SHD giant cell colony, pHG/pEJ co-transfer. Experiments were done as described in Table 2 legend. $\times 11$.

gEJ and pEJ transformants, respectively (P. L. Jack and R.F.N. unpublished data). We conclude from these data that a single copy of EJ c-Ha-ras1 can probably fully transform 4DH2, but that multiple copies may augment the transformation response.

In sharp contrast to the effect of pEJ on newly immortalized lines, no transformed foci were produced after transfection of pEJ into tertiary SHD cells. The only observable response was the presence of clear plaques in stained monolayers (mean of 10 dishes ($\pm$ s.d.): 144 ± 14) when large amounts of pEJ were transfected (Fig. 1e). On microscopic examination, plaques were seen to consist of clusters of 'giant' cells resembling SHD cells in the later stages of senescence; in a few cases plaques were associated with an area of more intensely stained (basophilic) giant cells at the periphery. We detected no immortal cells in isolates of 14 of the largest plaques and, because the cells rapidly entered senescence after the first subculture, we were unable to obtain sufficient cells for anchorage independence studies. Similar experiments with pEJ were performed on two early-passage lines of human dermal fibroblasts. No pEJ-induced foci were observed.

In view of the fact that a focus assay for transformation relies on extensive overgrowth of transformed clones on confluent monolayers during a relatively long incubation period post-transfection (Table 1 legend), we were concerned that a possible early effect of pEJ on the anchorage dependence of normal (pre-senescent) cells might not be detectable by this method, if the cells retain a limited capacity for proliferation. To test this possibility, we replated secondary SHD and Syrian hamster embryo (SHE) cells into soft agar immediately after pEJ transfection. SHE cells were included to represent a less differentiated population of normal cells, with greater division potential; the immortal line 4DH2 was used as a positive control. Cultures were observed daily for signs of colony formation. During the first 3–4 days, it appeared that a similar frequency ($\sim 1-2 \times 10^{-3}$) of anchorage-independent variants had been induced in all three types of cell by pEJ. However, the vast majority (>90%) of SHD and SHE soft agar colonies had a severely limited capacity for progressive growth thereafter (<50 cells per colony). Of those that attained a larger size (100–200 cells), 40 were isolated and returned to monolayer culture. All isolates

Table 2 Properties of Mtx^r colonies resulting from co-transfection of pHG and pEJ plasmids into tertiary cultures of Syrian hamster cells and the newly immortalized cell line, 4DH2

Recipient cell	DNA transfected	No. of colonies with non-transformed morphology/no. of dishes	No. of colonies with transformed morphology/no. of dishes	No. of giant cell colonies/no. of dishes
Syrian hamster dermal fibroblasts	a	0/10	0/10	0/10
	b	816/10	0/10	0/10
	c	139/10	0/10	125/10
Syrian hamster embryo cells	a	0/10	0/10	0/10
	b	872/10	0/10	0/10
	c	379/9	0/9	112/9
4DH2	a	0/10	0/10	0/10
	b	554/10	0/9	0/9
	c	78/9	567/9	0/9

Tertiary cultures of SHD and Syrian hamster embryo (SHE) cells together with the immortal cell line 4DH2 were transfected with pHG and pEJ DNA essentially as described in Table 1 legend. Replicate cultures (3.5×10^6 cells in 40 ml medium) of each cell type received calcium phosphate co-precipitates of either: a, 60 µg carrier (SHD) DNA alone; b, 45 µg carrier + 15 µg pHG DNA; or c, 30 µg carrier + 15 µg pHG DNA + 15 µg pEJ DNA, in 3 ml HEPES-buffered saline. Plasmid pHG, developed by O'Hare *et al.*¹⁵, contains genes capable of conferring selectable phenotypes on both bacteria (ampicillin resistance) and mammalian cells (methotrexate resistance). The latter property is due to the presence of a gene of prokaryote origin (*dhfr*) which codes for DHFR with low binding affinity for methotrexate. Other essential features of pHG are: presence of the simian virus 40 early promoter region (controlling *dhfr* transcription in mammalian cells) and rabbit β-globin donor and acceptor splice sites plus a polyadenylation site (essential for production of a functional *dhfr* transcript); for further details on the construction and properties of this plasmid see ref. 15. After an adsorption period of 18 h, the cultures were incubated in fresh medium for a further 24 h and then re-plated at a density of 10^5 cells per 9 cm dish in medium containing 10% FCS and 0.8 µM methotrexate (Sigma); this selection medium was renewed after 48 h. Methotrexate-resistant (Mtx^r) colonies were fixed and stained after an additional 7 days of incubation. Colonies were enumerated and classified on the basis of morphology (see Fig. 1) using a low-power dissecting microscope. To facilitate scoring, experiments were designed such that transfection produced ~50–100 Mtx^r colonies per 9 cm dish.

underwent senescence within three passages. In contrast, all pEJ-induced 4DH2 agar colonies grew progressively and produced continuous, fully transformed cell lines. Thus, finite cellular lifespan limits the transformation of normal cells by EJ c-Ha-ras1.

DNA-mediated gene transfer in mass cultures of mammalian cells is a relatively inefficient process, even with purified genes. However, one may substantially increase the chances that a non-selectable gene will be present in a particular clone of transfected cells by co-transfecting together with a suitable selectable marker^{7,14}. We adopted this approach in the present study to ensure successful uptake of pEJ into SHD cells. Thus, pEJ was co-transfected with the plasmid pHG¹⁵, which contains a dihydrofolate reductase (*dhfr*) gene of prokaryote origin. Because the encoded DHFR enzyme has a low binding affinity for methotrexate (Mtx), this gene will impart a methotrexate-resistant (Mtx^r) phenotype on mammalian cells.

Co-transfection experiments using pEJ and pHG were done with 4DH2, SHD and SHE cells. Transfection using pHG alone (Table 2 legend) produced Mtx^r colonies having normal morphology (Fig. 2a,c); frequencies were approximately 35% higher with SHD and SHE cells than with 4DH2 (Table 2). Co-transfection of the 4DH2 line with equal amounts of pHG and pEJ produced a similar number of colonies to that obtained with pHG alone. Of these colonies, 88% were morphologically transformed (Fig. 2b) and the cells had the same properties as pEJ transformants induced in mass culture. However, with SHD and SHE cells, co-transfection produced a substantial reduction in the total numbers of Mtx^r colonies recovered. Of those that did arise, a large proportion consisted of clusters of giant cells (Fig. 2d), resembling those in plaques in mass culture experiments. (The small number of cells in these clusters precluded the isolation of sufficient DNA, even from pooled isolates, to probe for EJ c-Ha-ras1 uptake, by Southern analysis.) No cells with altered lifespan or anchorage-independent properties were obtained from isolated individual Mtx^r colonies or pooled colonies from four 9-cm plates, cultured in the presence or absence of methotrexate.

Recent work which identified the activated *ras* transforming genes in some tumour cells and tumour cell lines (see also refs 16–19) unquestionably represents a substantial advance in our understanding of fundamental cancer mechanisms. However, it is as yet unclear exactly what role these activated genes have

in the multi-stage process by which most human and experimental cancers arise. NIH 3T3, which has to date been the only *in vitro* target cell used routinely for detecting such oncogenes, cannot be regarded as an ideal model for normality because it has previously undergone immortalization, is markedly heteroploid and will spontaneously generate malignant variants at a relatively high frequency. It is tempting, therefore, to speculate that *ras* gene activation represents one of a number of possible late steps in carcinogenesis, and that a preceding step (or steps) is required which is the key alteration providing the raw material for subsequent clonal evolution and progression. In our view carcinogen-induced immortality, which *in vitro* predisposes both fibroblasts¹² and epithelial cells²⁰ to further progression to malignancy, would seem to be a strong candidate for this critical early event. In the present study, we have obtained additional evidence in support of our hypothesis by showing that normal, pre-senescent cells are refractory to transformation by EJ c-Ha-ras1 but that such resistance completely disappears after immortalization by chemical and physical carcinogens.

This work was supported by grants from the Cancer Research Campaign and the MRC (G80/1084) and US NCI (RO1-CA-25807). We thank Professor R. A. Weinberg for providing pEJ, Dr K. O'Hare for pHG, Dr P. L. Jack for Southern analysis of DNA from transfectants and Mrs M. White for technical assistance.

Received 3 June; accepted 13 July 1983.

- Cooper, G. M. *Science* **218**, 801–806 (1982).
- Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404–409 (1982).
- Shih, C. & Weinberg, R. A. *Cell* **29**, 161–169 (1982).
- Pulciani, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 2845–2849 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
- Santos, E. *et al. Nature* **298**, 343–347 (1982).
- Tabin, C. J. *et al. Nature* **300**, 143–149 (1982).
- Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
- Taparowski, E. *et al. Nature* **300**, 762–765 (1982).
- Capon, D. J. *et al. Nature* **302**, 33–37 (1983).
- Farber, E. & Cameron, R. *Adv. Cancer Res.* **31**, 125–226 (1980).
- Newbold, R. F., Overell, R. W. & Connell, J. R. *Nature* **299**, 633–635 (1982).
- Wigler, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1373–1376 (1979).
- Wigler, M. *et al. Cell* **16**, 777–785 (1979).
- O'Hare, K., Benoist, C. & Breathnach, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1527–1531 (1981).
- Pulciani, S. *et al. Nature* **300**, 539–542 (1982).
- McCoy, M. S. *et al. Nature* **302**, 79–81 (1983).
- Balmain, A. & Pragnell, I. B. *Nature* **303**, 72–74 (1983).
- Hall, A., Marshall, C. J., Spurr, N. K. & Weiss, R. A. *Nature* **303**, 396–400 (1983).
- Pai, S. B., Steele, V. E. & Nettesheim, P. *Carcinogenesis* **4**, 369–374 (1983).
- Smith, B. L., Anisowicz, A., Chodosh, L. A. & Sager, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1964–1968 (1982).

DNA sequences of two yeast promoter-up mutants

David W. Russell* & Michael Smith

Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

David Cox, Valerie M. Williamson* & Elton T. Young

Department of Biochemistry, University of Washington, Seattle, Washington 98195, USA

The yeast *Saccharomyces cerevisiae* has three genetic loci encoding different alcohol dehydrogenase (ADH) isozymes: *ADC1*, which encodes the classical fermentative isozyme ADH^I; *ADR2*, which encodes the glucose-repressed isozyme ADH^{II}; and *ADM*, which encodes an ADH isozyme found associated with mitochondria. When yeast are grown on glucose, the *ADC1* gene is expressed, and the *ADR2* gene repressed^{1,3,4}. Conversely, growth on a non-fermentable carbon source such as ethanol or glycerol results in derepression of *ADR2*, and repression of *ADC1*. The *ADC1* and *ADR2* genes have been cloned and sequenced⁵⁻⁸, and a number of *cis*-acting mutations identified that cause constitutive expression of *ADR2*^{9,10}, and seem to fall into two classes¹¹. The most abundant class consists of mutants that cannot be fully derepressed, and do not revert to wild type at a detectable level: these are caused by the insertion of a transposable element into the 5'-flanking region of the gene^{6,12}. The second class of mutants do revert to a glucose-repressed phenotype at a detectable frequency, and when grown on non-fermentable carbon sources derepress *ADR2* to levels up to five times those found in wild-type cells¹⁰. We report here the sequencing of the 5'-flanking regions of two such promoter-up, constitutive *ADR2* mutants, in both of which the mutant phenotype is associated with an increase in length of a poly(A)·poly(T) tract 222 base pairs (bp) upstream of the gene.

The strategy used to clone the constitutively expressed *ADR2* gene from the *ADR3-4^c* and *ADR3-5^c* mutants is outlined in Fig. 1. Yeast genomic DNA was isolated from the appropriate strain, digested with *Bam*HI, and size-fractionated by agarose gel electrophoresis. Southern blot analysis of the digested DNA using the cloned wild-type *ADR2* gene as a probe did not reveal any gross rearrangement of the mutant *ADR2* genes (data not shown). Fragments approximately 7 kilobases (kb) long were isolated from preparative gels and cloned in the yeast-*Escherichia coli* shuttle vector YRp7 (ref. 13). Plasmids containing the *ADR2* gene were identified by colony hybridization¹⁴ and subsequently tested for their ability to complement yeast cells deficient in ADH activity⁵. For both mutants, the DNA used in the sequence analysis conferred constitutive *ADR2* expression on the transformed yeast.

Digestion with *Sau*3AI of the 1.9-kb *Bam*HI-*Hind*III fragment of the wild-type *ADR2* gene that contains the 5'-flanking region and part of the coding sequence (Fig. 1) yielded six fragments, of sizes 568, 509, 385, 234, 160 and 100 bp⁸. Digestion of the equivalent fragment of both mutant genes yielded the same five smaller fragments with *Sau*3AI, but the largest fragment (which contains the amino-terminal coding sequence of *ADR2* and 317 bp of 5'-flanking sequence) migrated with a length of about 640 bp. Moreover, this fragment migrated in a diffuse manner on 5% polyacrylamide gels when compared with co-electrophoresed fragments of both higher and lower molecular weights. Thus, both mutants seemed to

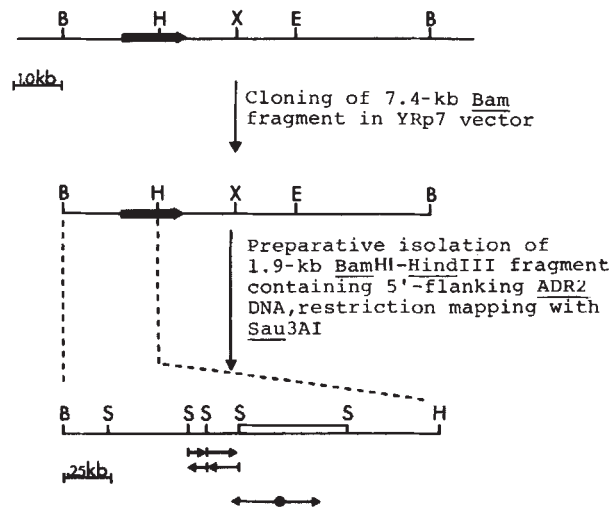


Fig. 1 Cloning of *ADR3-4^c* and *ADR3-5^c* DNA. The upper line represents a restriction endonuclease map of yeast genomic DNA in the *ADR2* region. Yeast DNA was isolated as described⁶, digested with *Bam*HI, and size-fractionated on 0.85% agarose gels. Fragments of ~7 kb were isolated by electroelution and ligated into *Bam*HI-cleaved YRp7 (ref. 13). *E. coli* RR1 cells were transformed to ampicillin resistance with the ligated DNA and screened by colony hybridization using the wild-type *ADR2* gene as a probe. Positive clones were characterized by restriction mapping and by complementation of function of yeast deficient in ADH activity. Restriction enzyme cleavage sites for *Bam*HI (B), *Hind*III (H), *Xho*I (X) and *Eco*RI (E) are indicated. The bold arrow indicates the location and direction of transcription of the *ADR2* gene. The 5'-flanking DNA of the mutant *ADR2* genes was further characterized by restriction mapping with *Sau*3AI (S) and by DNA sequence analysis (lower line). The open box indicates the 640-bp *Sau*3AI fragment having an altered electrophoretic mobility and discussed in the text. Arrows immediately below the bottom line indicate fragments cloned in M13mp7 (ref. 15) and used in DNA sequence analysis. Arrows diverging from the dot indicate DNA sequenced by the procedures of Maxam and Gilbert¹⁷.

have an insertion of approximately 70 bp near the 5' end of the gene.

The strategy we used to sequence the promoter regions of the two constitutive mutants is shown in Fig. 1. The 100- and 160-bp *Sau*3AI fragments were isolated and cloned in both orientations in the *Bam*HI site of the bacteriophage M13mp7 vector¹⁵. Sequence analysis of these clones using a universal primer¹⁵ and the terminator method¹⁶ did not reveal any differences between the mutant and wild-type DNAs. In a similar manner, the 640-bp *Sau*3AI fragment discussed above was isolated and cloned in M13. Sequence analysis of two such clones indicated, however, that this yeast DNA fragment was unstable in the mp7 vector and could not be recovered intact. Thus, the base-specific chemical degradation method¹⁷ was used to sequence the 640-bp *Sau*3AI fragments isolated from the original plasmids (Fig. 2). Both mutants have alterations in their DNAs adjacent to and within a block of 20 contiguous A·T bp which occurs 222 bp 5' to the coding sequence of the wild-type *ADR2* gene. A single G·C bp has been deleted 5' to the poly(A)·poly(T) tract in the *ADR3-5^c* DNA, 33 A·T bp have been inserted to make a total of 53 consecutive A·T bp, and 3 bp have been deleted 3' to the poly(A)·poly(T) tract (Fig. 2a). Similarly, in the *ADR3-4^c* DNA, a T·A bp has been substituted for the wild-type G·C bp 5' to the poly(A)·poly(T) tract, 34 additional A·T bp have been inserted within the poly(A)·poly(T) tract, and 3 bp have been deleted 3' to the poly(A)·poly(T) tract (Fig. 2b). Except for these changes (see Fig. 4a for a summary), no other differences were found in the 5'-flanking DNA of the mutant *ADR2* genes.

The correlation between the presence of the extended poly(A)·poly(T) tracts and adjacent changes and the similar phenotypes in the two mutants suggested that the effect on

* Present addresses: Department of Molecular Genetics, University of Texas Health Science Center at Dallas, 5323 Harry Hines Blvd, Dallas, Texas 75235, USA (D.W.R.); ARCO Plant Cell Research Institute, Dublin, California 94568, USA (V.M.W.).

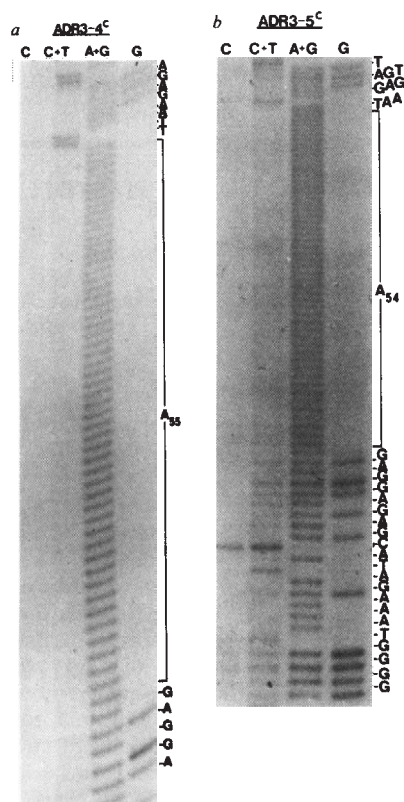


Fig. 2 DNA sequence analysis of *ADR3-4^c* and *ADR3-5^c*. The 1.9-kb *Bam*HI-*Hind*III fragment described in Fig. 1 was simultaneously digested with the restriction enzymes *Taq*I and *Rsa*I. *Taq*I sites were subsequently labelled at their 3' ends with [α -³²P]dCTP and the Klenow fragment of *E. coli* DNA polymerase I as described⁸. The appropriate fragment was sequenced by the methods of Maxam and Gilbert¹⁷. Electrophoresis was performed on thin gels (17×85 cm) as described by Sanger and Coulson²⁴. *a*, Autoradiogram of sequence derived from *ADR3-4^c* DNA in which 55 consecutive A residues are found. *b*, Autoradiogram of sequence derived from *ADR3-5^c* DNA in which 54 consecutive A residues occur. Sequence differences between the mutant and wild-type DNAs are summarized in Fig. 4a.

ADR2 gene expression was caused by these alterations. To ensure that mutations further 5' to the sequenced *Sau*3AI fragments were not responsible for the constitutive phenotype, plasmids containing the *Bam*HI-*Xho*I fragment from either the wild-type or *ADR3-4^c* clones were inserted into a yeast vector containing the loci *CEN3* and *TRP1*¹⁸. These plasmids have only two *Bcl*I sites, both within the yeast insert. The 5' *Bcl*I site overlaps the *Sau*3AI site that was the 5' limit of the DNA sequence analysis (Fig. 1). The other *Bcl*I site is 500 bp 3' to the *ADR2* structural gene⁸. The *ADR2*-containing *Bcl*I fragments from wild-type and *ADR3-4^c* DNA were exchanged, allowing the effect on *ADR2* activity of the flanking DNA 5' to the sequenced region to be determined. As shown in Fig. 3, the constitutive phenotype was associated with the *ADR3-4^c* *Bcl*I fragment containing the extended poly(A)-poly(T) tract and associated alterations, and not with upstream sequences, thus confirming that these changes are responsible for the constitutivity of *ADR3-4^c*.

The DNA sequencing results summarized in Fig. 4a are best interpreted in the light of a schematic representation of the wild-type *ADR2* gene (Fig. 4b). Previously, five *ADR2* mutants were characterized in which constitutive expression of *ADHII* was associated with the insertion of a yeast transposable element into the 5'-flanking region of the gene¹². These insertions occurred at positions -212, -189, -169, -161 and -125 relative to the *ADR2* coding sequence. One explanation for the loss of glucose repression in these transposable element mutants is that insertion occurred between regulatory sequences (*ADR2* distal) and promoter sequences (*ADR2* proximal). An examin-

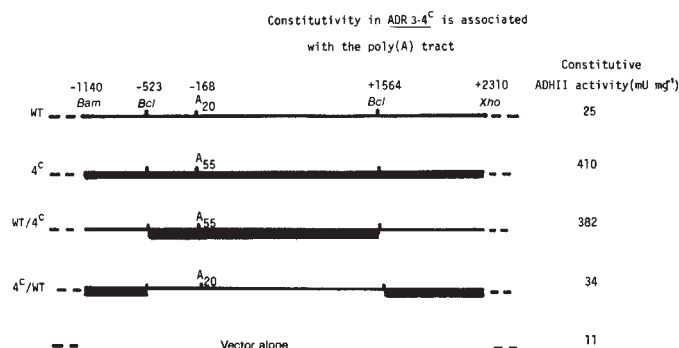


Fig. 3 Constitutivity in *ADR3-4^c* is associated with the extended A tract. The *Bam*HI-*Xho*I fragment of *ADR2* or *ADR3-4^c* was ligated into vector pBC3T1 (ref. 19) previously cut with *Bam*HI and *Sal*I. Transformants of a *Dam*⁻ strain of *E. coli* were selected on ampicillin. Appropriate transformants were grown and plasmid DNA was prepared and cut with *Bcl*I. The 2-kb *Bcl*I fragments from the wild-type and mutant clones were ligated into *Bcl*I-cut plasmids containing the mutant and wild-type flanking sequences, respectively. The plasmids were isolated and purified from transformed *E. coli* RR1 and used to transform *S. cerevisiae* strain 302-21 (α *trp1 his4 adc1-11, adr2-43, adm3*) to *Trp*⁺. Four yeast transformants for each plasmid were grown on *Trp*⁻ minimal medium containing 8% glucose and extracts were prepared and assayed for ADH activity. *ADR2* sequences, —; *ADR3-4^c* sequences, —; vector sequences, ---.

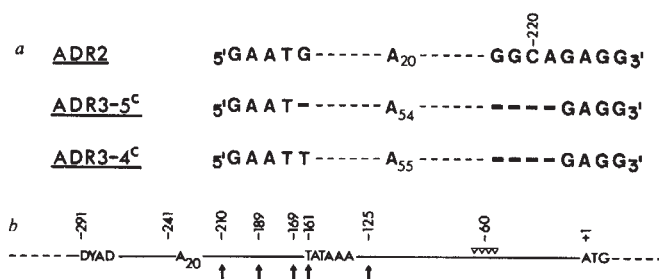


Fig. 4 *a*, Summary of the mutant and wild-type DNA sequences. -220 is the distance in nucleotide pairs from the coding region of *ADR2*. *b*, Organization of the 5'-flanking region of the wild-type gene. DYAD refers to a 22-bp sequence of perfect dyad symmetry. A₂₀ refers to a sequence of 20 consecutive A·T base pairs¹². Arrows indicate the sites of transposable element insertions in previously characterized *ADR2* constitutive mutants¹². Open triangles indicate the four sites of transcript initiation located ~60 bp 5' to the coding sequence of the gene⁸. ATG indicates the position of the N-terminal methionine-encoding DNA.

ation of sequences 5' to the leftmost transposable element insertion site (-212) revealed two features that were postulated to have a role in regulation. The first was a block of 20 consecutive adenine residues beginning at -241, and the second was a 22-bp sequence of perfect dyad symmetry located at -291 (Fig. 4b). We have shown here that constitutive and enhanced expression of *ADR2* may be brought about by an alteration in the A₂₀ sequence. In two different isolates, this alteration involves an addition of 33 or 34 adenine residues, together with the substitution and deletion of several base pairs flanking the poly(A) sequence (Fig. 4a).

The anomalous electrophoretic migration of the *Sau*3AI fragment containing the expanded poly(A)-poly(T) sequence, and the ability of these constitutive mutants to revert to a glucose-repressed phenotype may both be related to this unusual sequence. The presence of 54 or 55 consecutive adenine residues may increase the frequency of localized denaturation or breathing of the DNA within this sequence, giving rise to a more diffuse band during electrophoresis. In turn, this property of the DNA may facilitate recombination by increasing the frequency with which base-paired structures occur. The highly recombinogenic nature of these sequences is indicated by their

instability during propagation in the M13mp7 vector and, by analogy, to the reversion to a glucose-repressed phenotype characteristic of these constitutive mutants (ref. 11 and V.M.W., unpublished observations).

The mechanism by which these mutations arose is not clear. Although both mutants have an almost identical sequence, they represent independent isolates obtained after UV or ethyl methane sulphonate mutagenesis¹⁰. One possible source of the sequences is suggested by the finding of poly(A)·poly(T) tracts of variable length scattered throughout the yeast genome (ref. 19 and T. Petes, personal communication). The short A₂₀ duplex sequence upstream from *ADR2* might have been expanded in the mutant by recombination with such a tract. This event could be facilitated by sequence homologies flanking the adenine blocks. Unequal recombination between the *ADR2* poly(A)·poly(T) tracts during mitosis could also explain the expansion of these sequences. Alternatively, both mutations may have arisen during DNA replication, perhaps brought about by slippage of the DNA polymerase enzyme within the poly(A)·poly(T) sequence.

It is difficult to explain the loss of glucose repression in the *ADR3-4^c* and *ADR3-5^c* mutants. One possibility is that the upstream displacement of 29 or 31 bp of the putative glucose regulatory sequences occurring in these mutants is sufficient to render expression of the gene constitutive. However, these sequences have been shown to exert the appropriate regulatory effect when positioned as far as 600 bp 5' to the coding sequence of the *ADC1* gene¹⁸. Alternatively, the poly(A)·poly(T) tracts could function as a glucose-independent promoter. However, S₁ nuclease mapping has indicated that the 5' ends of the mRNA transcribed from the *ADR3-5^c* gene are identical to those found in wild-type *ADR2* cells (V.M.W., unpublished observations). A more intriguing explanation for constitutive and enhanced expression may be related to chromatin structure. Recent studies have indicated that poly(A)·poly(T) sequences as short as 20 bp are disfavoured during nucleosome formation, and that an 80-bp sequence is essentially excluded from this process^{20,21}. A disruption of the chromatin structure brought about by 54 or 55 consecutive A·T base pairs might lead to constitutive expression of the *ADR2* gene in the mutants analysed here. This same disruption could facilitate the entry of the RNA polymerase enzyme, leading to enhanced expression of the gene. Chromatin conformational changes have been shown to occur in the 5'-flanking region of the *ADR2* gene following derepression²², and recent evidence indicates that chromatin structure may have a role in regulating the expression of the *HMR* and *HML* loci of the yeast mating-type switching system²³.

We thank T. Petes, H. G. Martinson, K. Nasmyth, M. Ciriacy and K. E. Van Holde for communicating results before publication and for helpful discussions. D.W.R. was a Damon Runyon-Walter Winchell Fellow. M.S. is a career investigator of the MRC, Canada. This work was supported by an MRC grant to M.S., and by NIH grant GM 26079 awarded to E.T.Y.

Received 25 May; accepted 10 June 1983.

1. Ciriacy, M. *Mutat. Res.* **29**, 315-326 (1975).
2. Ciriacy, M. *Molec. gen. Genet.* **138**, 157-164 (1975).
3. Luttorf, U. & Megnet, R. *Archs Biochem. Biophys.* **126**, 933-944 (1968).
4. Denis, C., Ciriacy, M. & Young, E. T. *J. molec. Biol.* **148**, 355-368 (1981).
5. Williamson, V. M., Bennetzen, J., Young, E. T., Nasmyth, K. & Hall, B. D. *Nature* **283**, 214-216 (1980).
6. Williamson, V. M., Young, E. T. & Ciriacy, M. *Cell* **23**, 605-614 (1981).
7. Bennetzen, J. L. & Hall, B. D. *J. biol. Chem.* **257**, 3018-3025 (1982).
8. Russell, D. W., Smith, M., Williamson, V. M. & Young, E. T. *J. biol. Chem.* **258**, 2674-2682 (1983).
9. Ciriacy, M. *Molec. gen. Genet.* **145**, 327-333 (1976).
10. Ciriacy, M. *Molec. gen. Genet.* **176**, 427-431 (1979).
11. Ciriacy, M. & Williamson, V. M. *Molec. gen. Genet.* **182**, 159-163 (1981).
12. Williamson, V. M., Cox, D., Young, E. T., Russell, D. W. & Smith, M. *Molec. cell. Biol.* **3**, 20-31 (1983).
13. Struhl, K., Stinchcomb, D. T., Scherer, S. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1035-1039 (1979).
14. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
15. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9**, 309-321 (1981).
16. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
17. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
18. Beier, D. R. & Young, E. T. *Nature* **300**, 724-728 (1982).

19. Walmsley, R. M., Szostak, J. W. & Petes, T. D. *Nature* **302**, 84-86 (1983).
20. Kunkel, G. R. & Martinson, H. G. *Nucleic Acids Res.* **9**, 6869-6888 (1981).
21. Prunell, A. *EMBO J.* **1**, 173-179 (1982).
22. Sledziewski, A. & Young, E. T. *Proc. natn. Acad. Sci. U.S.A.* **79**, 253-256 (1982).
23. Nasmyth, K. A. *Cell* **30**, 567-578 (1982).
24. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107-110 (1978).

Interior turns in globular proteins

George D. Rose & William B. Young

Department of Biological Chemistry, Hershey Medical Center, Pennsylvania State University, Hershey, Pennsylvania 17033, USA

Lila M. Gierasch

Department of Chemistry, University of Delaware, Newark, Delaware 19711, USA

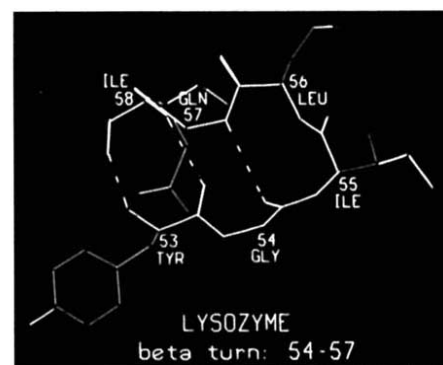
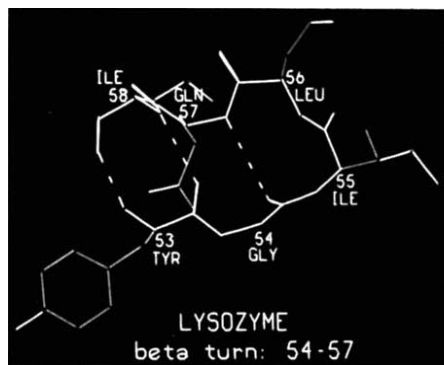
Reverse turns are specific sites in proteins at which the polypeptide chain changes its overall direction¹⁻⁶. This category of secondary structure enables the chain to turn a corner, and its frequent occurrence^{7,8} is the geometric basis for the ultimate globular shape of the protein. β -Turns in particular are comprised of four consecutive residues with a stereochemistry⁹ that constrains the turn to be polar. In consequence, turns are almost always situated at the surface of the protein, in contact with solvent water. We have searched proteins of known structure and find that, on occasion, a turn may be buried within the hydrophobic interior of the molecule. In every instance of a buried turn, one or more solvent molecules were also found in a hydrogen-bonded complex with main-chain atoms of the turn residues. These bound water molecules appear to function as an integral part of the protein structure.

β -Turns, as defined by Venkatachalam⁹, consist of four consecutive residues with a hydrogen-bond joining the C=O of the first turn residue (*i*) to the N-H of the last residue (*i* + 3); this steric arrangement requires that any hydrogen bond involving the remaining two residues (*i* + 2, *i* + 3) be formed with partners outside the turn. An isolated β -turn is already highly polar due to these unpaired N-H and C=O groups. In the context of a protein, polarity is further enhanced because turn formation is promoted at chain sites where the average linear hydrophobicity is locally minimal^{10,11}. Thus, it is hardly surprising that turns are almost always situated at the molecular surface^{1,4,10-12}, in contact with solvent water.

The Brookhaven data base¹³ was screened for interior turns in refined structures that include solvent coordinates. Twenty-two proteins meeting these conditions were examined, and six with interior turns were found, including three homologous serine proteases (see Table 1). Each turn satisfied the following criteria: (1) The dihedral angles for the middle residues (*i* + 1, *i* + 2) are near their ideal values. Following Venkatachalam⁹, ideal values are taken to be (-60, -30) and (-90, 0) for type I and (-60, 120) and (80, 0) for type II. (2) The distance from the carbonyl oxygen of the first turn residue (*i*) to the amide nitrogen of the last turn residue (*i* + 3) is small enough to permit hydrogen-bonding (<3.4 Å). An interior turn in phospholipase was allowed as an exception because in this case the distance from C=O (*i* - 1) to N-H (*i* + 3) is 3.15 Å, forming a five-residue, hydrogen-bonded turn. (3) The residues were required to be at a locally minimal radius of curvature⁶, indicating the presence of a turn site. These are comparatively stringent criteria, as evidenced by the small C α (*i*) - C α (*i* + 3) distances which are well below the usual 7 Å cutoff^{2,8}.

A detailed atomic description of the protein surface was generated according to the method of Lee and Richards¹⁴. In this approach, the molecular surface is defined by those atoms that are accessible to a water-sized probe; remaining atoms comprise the interior. We have used an implementation of this method due to Connolly¹⁵ that is tailored for use with computer graphics.

Fig. 1 Stereoview of a β -turn in lysozyme, residues 54–57, shown in the context of a containing hexapeptide segment, residues 53–58. Broken lines depict hydrogen bonds, which connect C=O(54) to N–H(57) in the turn, as well as C=O(53) to N–H(58) and N–H(53) to C=O(58) in the hexapeptide.



For each turn, the fraction buried was determined by subtracting the solvent-accessible surface area from total surface area and then dividing by total surface area. Calculated values are listed in Table 1, columns 7–9. By way of illustration, the stippled surface in Fig. 2a depicts the solvent-accessible fraction of the lysozyme turn, residues 54–57, while Fig. 2b shows the total surface that results when these residues are viewed in isolation. Comparison of the two situations indicates that the turn is indeed buried, an interpretation confirmed by Table 1; 84% of the turn's surface is rendered inaccessible by the surrounding presence of the protein.

The proteins used in this study and their Brookhaven file names include phospholipase A₂ (1BP2)¹⁶, γ -chymotrypsin A (2GCH)¹⁷, lysozyme (6LYZ)¹⁸, β -trypsin (3PTP)¹⁹, staphylococcal nuclease (1SNS)²⁰ and trypsinogen (1TGB)²¹. Visualization was achieved with an Evans and Sutherland graphics system using the GRAMPS interpreter written by T.J. O'Donnell. Interatomic distances and angles were calculated directly from X-ray coordinates.

The turns under study are all type I and type II β -turns, and Table 1 lists pertinent information about each.

A striking finding is the presence of one or more tightly complexed water molecules in the vicinity of each turn. The water molecule associated with the lysozyme turn is illustrated in Fig. 2, while in Fig. 3 it is displayed with its hydrogen-bonded neighbours from both the turn and the rest of the protein.

Each of the water molecules bound to an interior turn has at least three hydrogen bonds. Moreover, each polar group in these turns is hydrogen bonded either to a water molecule or to another interior polar group of the protein. Instances of bifurcated hydrogen bonds to these groups were also observed. Table 2 lists water ligands together with all hydrogen-bonded partners closer than 3.4 Å.

To assess whether these complexed water molecules are themselves buried within the protein, their number density was calculated using all protein atoms within 8 Å of each solvent molecule. In Table 2, the pair of numbers in parentheses gives the range of number densities computed for all water molecules

Table 1 Interior turns

Residues in turn	Sequence	Dihedral angles	C=O...N distance	C $^{\alpha}(i)$ –C $^{\alpha}(i+3)$ distance	Atoms in turn	Accessible surface area	Total surface area	% Buried
1. γ -Chymotrypsin A (2GCH) 27–30	Trp-Pro-Trp-Gln	$i+1$: –59, –25 $i+2$: –81, –12 type I	2.94	5.61	44	19.39	447.33	95.7
2. γ -Chymotrypsin A (2GCH) 194–197	Asp-Ser-Gly-Gly	$i+1$: –53, 144 $i+2$: 83, –20 type II	3.12	5.76	22	12.34	268.92	95.4
3. β -Trypsin, diisopropylphosphoryl inhibited (3PTP) 27–30	Val-Pro-Tyr-Gln	$i+1$: –68, –15 $i+2$: –91, –3 type I	3.14	5.80	35	74.19	368.08	79.8
4. β -Trypsin, diisopropylphosphoryl inhibited (3PTP) 194–197	Asp-Ser-Gly-Gly	$i+1$: –46, 142 $i+2$: 90, –13 type II	3.11	5.85	22	8.72	266.97	96.7
5. Trypsinogen (1TGB) 27–30	Val-Pro-Tyr-Gln	$i+1$: –53, –21 $i+1$: –99, 0 type I	3.22	5.74	35	68.04	350.61	80.6
6. Trypsinogen (1TGB) 194–197	Asp-Ser-Gly-Gly	$i+1$: –47, 145 $i+2$: 97, –15 type II	3.32	5.95	22	53.78	269.55	80.0
7. Bovine pancreas phospholipase A ₂ (1BP2) 26–29	Gly-Cys-Tyr-Cys	$i+1$: –90, –22 $i+2$: –111, –30 type I	4.93	6.07	28	29.96	369.94	91.9
8. Hen egg lysozyme (6LYZ) 54–57	Gly-Ile-Leu-Gln	$i+1$: –62, –31 $i+2$: –108, 7 type I	3.33	4.73	29	53.90	334.90	83.9
9. Staphylococcal nuclease (1SNS) 19–22	Asp-Gly-Asp-Thr	$i+1$: –4, –80 $i+2$: –86, 25 type II'	2.54	4.95	27	74.52	268.64	72.3

Table 2 Water hydrogen bonds

	Residue or water no.	Group	Distance	No. density (range)		Residue or water no.	Group	Distance	No. density (range)
1. γ -Chymotrypsin A (27-30)				(11-124)	5. Trypsinogen (27-30)			(16-122)	
66W	27	N-H	3.11	113	717W	27	N-H	2.80	103
	27	C=O	2.79			27	C=O	2.96	
	Water	121	2.77			25	N-H	3.29	
17W	28	C=O	2.78	111		26	N-H	3.29	
	30	C=O	2.86			Water	604	2.91	
	69	N-H	2.83			Water	709	3.19	
22W	30	C=O	3.10	100	716W	28	C=O	2.86	107
	67	C=O	2.85			30	C=O	2.94	
	70	N-H	3.06			69	N-H	2.86	
2. γ -Chymotrypsin A (194-197)				(11-124)	708W	30	C=O	2.61	110
9W	194	C=O	2.97	117		66	C=O	3.26	
	139	C=O	2.76			69	N-H	3.19	
	Water	14	2.88			70	N-H	2.79	
26W	196	C=O	2.76	108		Water	709	3.05	
	53	C=O	3.13		6. Trypsinogen (194-197)			(16-122)	
	45	O γ	2.98		401W	194	C=O	3.12	113
3. β -Trypsin (27-30)				(7-121)		139	C=O	3.23	
28W	27	N-H	2.79	103	406W	196	C=O	2.82	107
	27	C=O	2.96			45	N-H	3.12	
	26	N-H	3.14			45	O γ	2.51	
	25	N-H	3.30			53	C=O	2.97	
	Water	51	2.91		7. Phospholipase A ₂ (26-29)			(10-117)	
	Water	52	3.15		22W	26	C=O	3.07	96
42W	28	C=O	2.89	113		29	C=O	3.00	
	30	C=O	2.73			120	N ζ	2.64	
	69	N-H	2.79			Water	73	2.87	
41W	30	C=O	2.88	107	26W	26	N-H	3.30	95
	65	C=O	2.65			26	C=O	2.87	
	69	N-H	3.32			24	C=O	3.19	
	70	N-H	2.96			117	N-H	3.34	
52W	27	C=O	3.35	113		118	N-H	2.72	
	30	C=O	3.19		5W	28	C=O	3.13	95
	70	C=O	2.96			30	N-H	2.99	
	Water	28	3.15			30	C=O	3.08	
	Water	46	2.82			49	O δ 1	3.22	
4. β -Trypsin (194-197)				(7-121)		Water	6	2.56	
37W	194	C=O	2.87	117	8. Lysozyme (54-57)			(10-106)	
	139	C=O	2.85		1W	56	N-H	3.09	106
	30	O ϵ 1	3.33			53	C=O	2.94	
	Water	16	2.86			91	O γ	3.25	
69W	196	C=O	2.83	108	9. Staphylococcal nuclease (19-22)			(46-122)	
	197	C=O	3.38		11W	20	N-H	2.57	108
	45	O γ	2.50			19	O γ 1	3.04	
	53	C=O	3.37			41	O γ 1	2.53	
						42	C=O	2.56	

The table lists all water ligands for each buried turn together with other possible hydrogen-bonding partners closer than 3.4 Å. Water molecules are of the form nW , where n is their Brookhaven index number¹³. The number density, shown in the far right column, is the calculated number of protein atoms contained within an 8 Å sphere centred at the given water molecule. A high number density indicates that the water is well buried within the protein¹¹. A single table entry for each protein, shown in parentheses, gives the range of number densities observed for all the water molecules in that protein. Waters associated with buried turns are seen to be at the high end of this range in each case.

in each protein. The table also gives the number density of individual water molecules that are hydrogen bonded to an internal turn. It is apparent that these latter values are at the high end of the range and resemble corresponding values found for the most deeply buried α -carbons in globular proteins¹¹. Thus, water molecules in complex with interior turns are themselves deeply buried within the protein and not part of the bulk solvent.

The loci of buried waters are known to be conserved in related proteins^{22,23}. The present study includes two internal turns within each of three serine proteases. Both turns are found in regions of conserved sequence in the serine proteases²⁴. The disposition of the water molecules in complex with these turns exhibits a similarly conserved pattern, as shown in Tables

1 and 2, with five buried water molecules hydrogen-bonded to most of the same donors and acceptors in each protein.

Groups from many of these buried turns assume functional roles in their respective proteins. Residues 194-197 in the serine proteases contain active site Ser 195 (ref. 25). In phospholipase, the entire turn is part of the calcium binding loop¹⁶, and in the staphylococcal nuclease turn, calcium is bound to O δ 1 of Asp 21 (ref. 20). Finally, the lysozyme turn and its bound water impinge on that enzyme's active site²⁶.

A turn situated in a hydrophobic environment such as the interior of a protein must find a way of neutralizing its polar groups without recourse to bulk solvent. In addition, hydrogen-bonding partners for the peptide groups of the middle residues must be found external to the turn itself, for reasons mentioned

Fig. 2 *a*, Stereoview of the lysozyme β -turn, residues 54–57, showing the solvent-accessible surface. Solvent access to the segment is calculated with a program by Connolly¹⁵ and is depicted by the stippled surface. The turn is seen to be buried within the interior of the protein: only 5 of the 29 atoms in turn proper (residues 54–57) can establish even limited contact with solvent, that is, C γ ¹ and C γ ² in Ile 55 and C γ , C δ ¹ and C δ ² in Leu 56. The position of a tightly complexed water molecule, visible in the X-ray map, is indicated by a W. *b*, Stereoview of the lysozyme turn, residues 54–57, treated as an isolated segment, with solvent-accessible regions depicted as stippled surface. In the absence of the surrounding protein, the entire turn can be solvated.

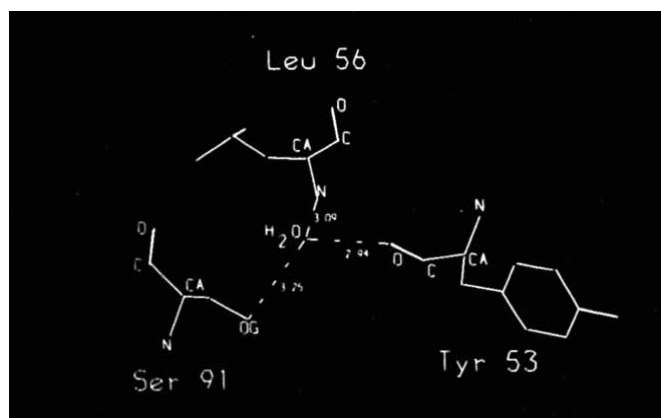
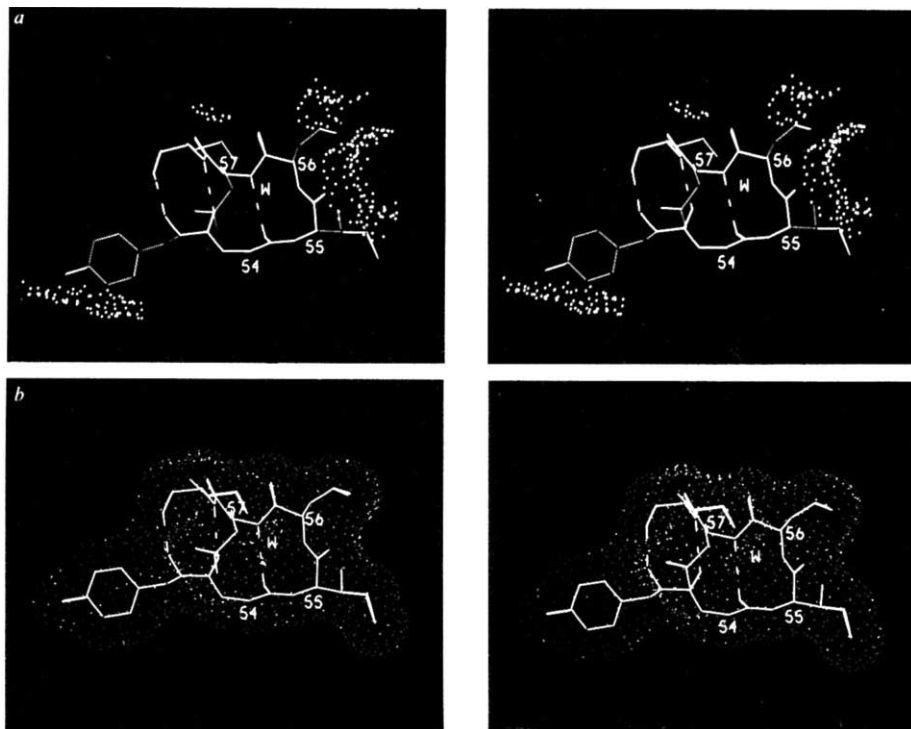


Fig. 3 A specific water molecule, visible in the X-ray map as HOH¹ (ref. 18), is illustrated together with its protein ligands. The water is hydrogen bonded to the C=O in Tyr 53, to the N—H in Leu 56, and to the O γ in Ser 91. These three residues are shown projected onto the O(53)—N(56)—O γ (91) plane; the water oxygen is 0.78 Å below the plane.

previously. These constraints have been satisfied for the turns in the present study by incorporating water molecules into the structure as though they were prosthetic groups. With this addition, buried chain segments are apparently able to adopt a favourable turn geometry in what would otherwise be an unfavourable environment.

The present findings represent a specific case of structural hydration, as defined by Kuntz and Kauzmann²⁷. Edsall and McKenzie²³ have also discussed these issues, and proposed four categories of water molecules in globular proteins. One of their categories is the internal water found within folded peptide chains²³.

The finding that interior turns are accompanied by bound water molecules may also have application to membrane proteins. The membrane resembles the protein's own interior to the extent that both are environments of low polarity as a consequence of close-packed alkyl groups and sparing water. Further, the fact that proteins do execute turns within the membrane can be inferred from cases where a protein penetrates the lipid bilayer and re-emerges on the same side without completely traversing the membrane^{28,29}.

Because peptide chain turns are constrained to be polar by their geometry, a turn within an apolar lipid is expected to be at least as thermodynamically unfavourable as the buried turns in the present study. It is possible that individual water molecules can contribute to the stability of a protein within a membrane by providing analogous site-specific polarity. Such water molecules would then constitute a reversible mechanism to modulate a membrane protein's transition from the polar cytoplasm to the apolar lipid.

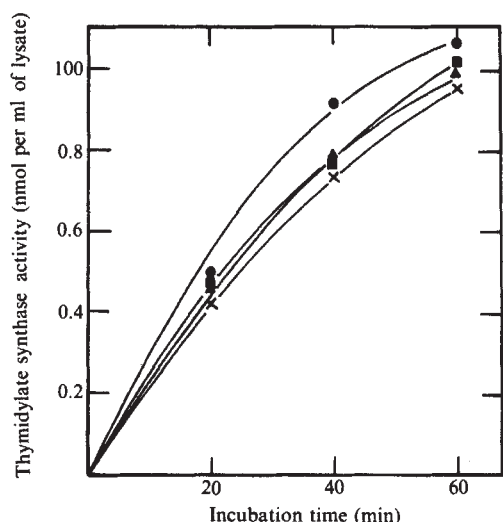
We thank Richard Lee, Tony Kossiakoff, Lyndon Hibbard and Robert Wood for useful advice. This work was supported by NIH grants GM 29458 (G.D.R.) and GM27616 (L.M.G.), a Research Career Development Award to G.D.R., a Research Service Award for short-term research training of medical students to W.B.Y., and grants from Merck, Sharpe & Dohme Laboratories and from Smith, Kline & French to L.M.G. This paper is dedicated to Walter Kauzmann on the occasion of his retirement.

Received 10 January; accepted 10 June 1983.

- Kuntz, I. D. *J. Am. chem. Soc.* **94**, 8568–8572 (1972).
- Lewis, P. N., Momany, F. A. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2293–2297 (1971).
- Crawford, J. L., Lipscomb, W. N. & Schellman, C. G. *Proc. natn. Acad. Sci. U.S.A.* **70**, 538–542 (1973).
- Levitt, M. & Chothia, C. *Nature* **261**, 552–558 (1976).
- Richardson, J. S. *Adv. Protein Chem.* **34**, 167–339 (1981).
- Rose, G. D. & Seltzer, J. P. *J. molec. Biol.* **113**, 153–164 (1977).
- Rose, G. D. & Wetlauffer, D. B. *Nature* **268**, 769–770 (1977).
- Chou, P. Y. & Fasman, G. D. *J. molec. Biol.* **115**, 135–175 (1977).
- Venkatachalam, C. M. *Biopolymers* **6**, 1425–1436 (1968).
- Rose, G. D. *Nature* **272**, 586–590 (1978).
- Rose, G. D. & Roy, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4643–4647 (1980).
- Smith, J. A. & Pease, L. G. *CRC crit. Rev. Biochem.* **8**, 315–399 (1980).
- Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535–542 (1977).
- Lee, B. K. & Richards, F. M. *J. molec. Biol.* **85**, 379–400 (1971).
- Connolly, M. *Quantum Chemistry Program Exchange Bull.*, Program 429, **75** (Indiana University Chemistry Department, Bloomington, 1981).
- Dijkstra, B. W., Kalk, K. H., Hol, W. G. J. & Drenth, J. *J. molec. Biol.* **147**, 97–123 (1981).
- Cohen, G. H., Silverton, E. W. & Davies, D. R. *J. molec. Biol.* **148**, 449–479 (1981).
- Diamond, R. *J. molec. Biol.* **82**, 371–391 (1974).
- Chambers, J. L. & Stroud, R. M. *Acta crystallogr. B* **35**, 1861–1874 (1979).
- Cotton, F. A., Hazen, E. E. Jr & Legg, M. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2551–2555 (1979).
- Fehlhammer, H., Bode, W. & Huber, R. *J. molec. Biol.* **111**, 415–438 (1977).
- Edsall, J. T. *Am. chem. Soc. Symp. Ser.* **127** (1980).
- Edsall, J. T. & McKenzie, H. A. *Adv. Biophys.* **16**, 53–183 (1983).
- Greer, J. *J. molec. Biol.* **153**, 1027–1042 (1981).
- Stroud, R. M., Kossiakoff, A. A. & Chambers, J. L. *A. Rev. Biophys. Bioengng* **6**, 177–193 (1977).
- Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C. & Rupley, J. A. *The Enzymes* **7**, 665 (1972).
- Kuntz, I. D. & Kauzmann, W. *Adv. Protein Chem.* **28**, 239–345 (1974).
- Dailey, H. A. & Strittmatter, P. *J. biol. Chem.* **256**, 3951–3955 (1981).
- Brunner, J. *et al. J. biol. Chem.* **254**, 1821–1828 (1979).

'Inhibitor evidence for allosteric interaction in the replisome multienzyme complex' by G. Prem veer Reddy and Arthur B. Pardee, *Nature* 304, 86-88 (1983).

By an unfortunate error in the editorial office, the figures accompanying this article on page 87 were wrongly printed. Fig. 1 was omitted and Fig. 2C was labelled as Fig. 1. The page as it should have appeared is reproduced below.



As shown in Fig. 1, hydroxyurea, novobiocin and aphidicolin had no inhibitory effects on thymidylate synthase activity in crude soluble extracts of S phase CHEF/18 cells as measured by the ^3H - H_2O formed from added $[\text{5-}^3\text{H}]\text{dUMP}$ on its conversion to dTMP^{5,6}. Thymidylate synthase, *per se*, is not sensitive to these inhibitors. In striking contrast, as shown in Fig. 2A-C, at similar or lower concentrations, they caused dramatic parallel inhibitions of thymidylate synthase activity and DNA synthesis measured in intact cells. These results demonstrate that thymidylate synthase activity is inhibited in intact cells by three agents which affect three other enzymes associated with DNA biosynthesis.

These observations could be accounted for if a metabolite such as dTTP accumulated during inhibition of DNA synthesis,

Fig. 1 Effect of hydroxyurea (x), novobiocin (Δ) and aphidicolin (■) on thymidylate synthase activity in soluble extracts of S phase CHEF/18 cells. ●, Control. CHEF/18 cells were grown and synchronized by isoleucine starvation as described elsewhere². When the cells were in mid-S phase, 12 h after release from an isoleucine block, they were trypsinized and suspended in ice-cold Dulbecco's modified Eagle's medium (DMEM). These cells were pelleted and resuspended in buffer A (mM): 150 sucrose, 80 KCl, 35 HEPES pH 7.4, 5 potassium phosphate pH 7.4, 5 MgCl₂, 0.5 CaCl₂, at a density of 5×10^7 cells ml⁻¹ and homogenized in a Potter-Elvehjem glass homogenizer by applying 10 strokes at setting 3 in a Sorvall Omni-mixer (type 17105, DuPont Instruments) with a teflon pestle. The nuclei were obtained by centrifugation at 850g for 10 min. The nuclear pellet was suspended in ice-cold buffer A to a density of $\sim 5 \times 10^7$ nuclear particles ml⁻¹, then sonicated in four pulses of 30 s each, using a Branson Sonifier (Model 185) at setting 2. The nuclear sonicate was centrifuged at 25,000 r.p.m. at 4 °C for 1 h in a Beckman SW 50.1 rotor with an adapter which allows the use of 0.8-ml tubes, and the supernatant was used as the soluble extract for thymidylate synthase assays, measured as described elsewhere¹. Inhibitors were included in the reaction mixture at the following concentrations: hydroxyurea, 0.5 mM; novobiocin, 0.5 mM; aphidicolin, 10 $\mu\text{g ml}^{-1}$.

and if this metabolite feedback-inhibited thymidylate synthase, or affected transport of ^3H -dUrd into the cell or kinase activity. The first condition seems unlikely because hydroxyurea prevents accumulation of deoxynucleotides⁷. The second condition appears to be ruled out because thymidylate synthase activity in permeabilized cells was not inhibited by excess dTMP or dTTP (50 μM) when the substrate, dUMP, was only 16 μM (ref. 1). Furthermore, hydroxyurea did not limit the accumulation of radioactive thymidine in acid-soluble pools of L929 cells⁷.

Thymidylate synthase activity is present throughout the cell cycle, but in intact or permeabilized cells it is high only when the cells are in S phase^{1,5}. During this phase thymidylate synthase is part of a sedimentable entity¹. From these observations

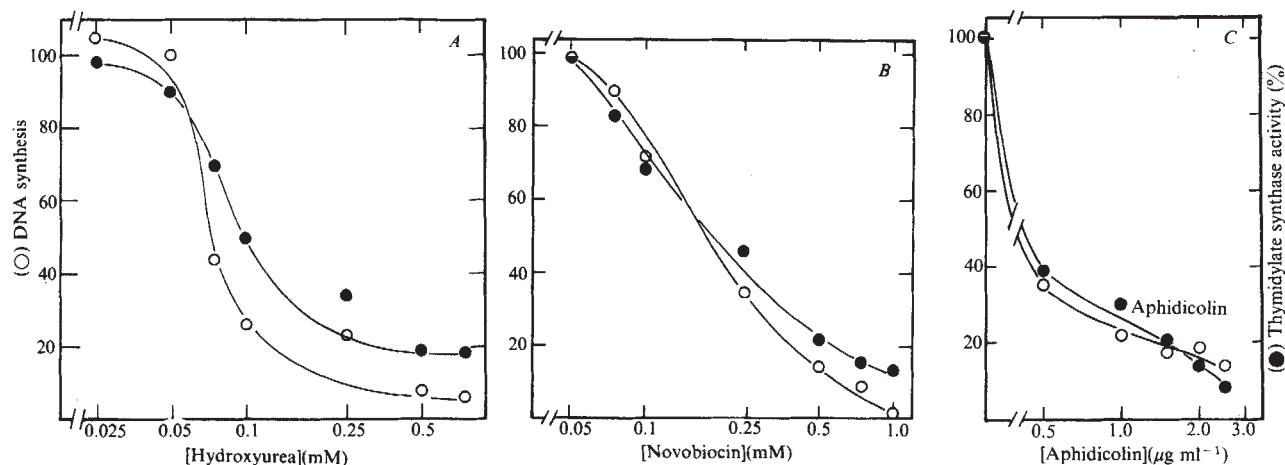


Fig. 2 Effect of inhibitors on DNA biosynthesis and thymidylate synthase activity in intact S phase CHEF/18 cells. **A**, Effect of hydroxyurea; **B**, effect of novobiocin; **C**, effect of aphidicolin. Open symbols, $[\text{6-}^3\text{H}]\text{dUrd}$ incorporation; solid symbols, ^3H - H_2O formed from $[\text{5-}^3\text{H}]\text{dUrd}$. **Methods:** CHEF/18 cells were synchronized as described in Fig. 1 legend. Duplicate culture dishes were pulsed for 1 h at 37 °C with undiluted 5 $\mu\text{Ci ml}^{-1}$ $[\text{6-}^3\text{H}]\text{dUrd}$ (24.5 Ci mmol⁻¹; NEN) or with 10 $\mu\text{Ci ml}^{-1}$ $[\text{5-}^3\text{H}]\text{dUrd}$ (13 Ci mmol⁻¹; Schwarz Mann). Cells pulsed with $[\text{6-}^3\text{H}]\text{dUrd}$ were processed to estimate the rate of DNA biosynthesis by washing monolayers twice with ice-cold phosphate-buffered saline (PBS), three times with 10% (w/v) trichloroacetic acid (TCA) and extracting the acid-insoluble material with 1 ml NaOH (0.2 M). Aliquots (100 μl) of alkaline extract were counted to determine moles of deoxyuridine incorporated into DNA. In control culture dishes, 100% $[\text{6-}^3\text{H}]\text{dUrd}$ incorporation into acid-insoluble material was equivalent to 6.9 ± 0.2 pmol per 10^6 cells per min. Culture dishes pulsed with $[\text{5-}^3\text{H}]\text{dUrd}$ were processed as follows to estimate thymidylate synthase activity in intact cells: the medium from these dishes was collected without disturbing the monolayer and analysed for ^3H - H_2O content formed by synthesis of dTMP from $[\text{5-}^3\text{H}]\text{dUMP}$, within the cell⁵. ^3H - H_2O was separated from $[\text{5-}^3\text{H}]\text{dUrd}$ by filtering the medium through a 1-cm charcoal (activated Norit) bed in a Pasteur pipette, which retained >99% of the $[\text{5-}^3\text{H}]\text{dUrd}$. The filtrate contained all the released tritium as ^3H - H_2O ; 100 μl of filtrate were counted for radioactivity. In control cultures 100% ^3H - H_2O formed from $[\text{5-}^3\text{H}]\text{dUrd}$ was calculated to be 6.0 ± 0.5 pmol per 10^6 cells per min. To test the effect of inhibitors, various concentrations of aphidicolin (Wako), hydroxyurea (Sigma) or novobiocin (Sigma) were added to the culture dishes 30 min before pulsing with radioactive nucleosides, and processed as described above.

Speciation in molluscs from Turkana Basin

SEVERAL of Williamson's conclusions on speciation in the freshwater molluscs of the Turkana Basin¹ are based on assumptions which need further examination. Here we comment on some relevant research in the Koobi Fora area.

Although the snail faunas collected by Williamson represent 'life and death assemblages'¹, he does not discuss the implications of this. We feel that the nature of the process of preservation may explain some morphological patterns in these fossil snails as well as the "instantaneous speciation" model proposed by Williamson. Our (A.C.) extensive bottom dredging of Lake Turkana shows that its modern molluscan assemblage consists of a mixture of stunted, transparent, living forms and long-dead, robust (bleached) specimens (with a few intermediate shells). The dead shells have been eroded from elevated Holocene and pre-Holocene lake beds. They occur both in shallow water (in which case most of the shells are broken) and in deeper water (where both broken and whole shells are present). The same patterns of deposition are found in Lake Naivasha (central Kenya rift valley).

Transported molluscan deposits in these lakes may therefore be mixed faunas containing individuals derived from sediments of different ages and of differing environmental origin. Assemblages of living snails in these modern lakes are much rarer than are such transported assemblages, and our observations have convinced us that the fossil deposits of Turkana Basin resemble those of the modern lake in this respect. We therefore challenge Williamson's statement² that "the faunal units concerned seem to be undisturbed life assemblages with no evidence of reworking" and note that in his original article¹ he states that they "represent both life and death assemblages". Given both the lack of information on the ancient environment and the nature of the evidence from the modern lake, we feel that Williamson's morphologically deviant populations (at the Suregei Complex, Lower Member and Guomde levels) may not represent instantaneous speciation as he suggests; if the deposits sampled at these horizons consist of transported assemblages the faunas could represent a mixture of individuals from geologically different periods or from geographically separated environments. Morphological deviants from one horizon would then reflect morphological change over extended periods of time or over

long distances, and could say little about the nature of morphological changes in the single fossil populations.

Williamson bases his argument for rapid speciation on his observations of rapid morphological change at three stratigraphic levels. However, we believe that 'rapid' morphological shifts are not well substantiated for at least two of these. He documents his lineages using evidence from *Bellamyia unicolor*. A summary of a canonical variate analysis for this lineage (p. 440 of ref. 1) illustrates increased phenotypic deviance at three levels, between populations 8–12c (Suregei Complex), 27–29 (Lower Member) and at 88 (Guomde Formation). There are several potential problems of interpretation here. In particular, Williamson's assertion that the periods of character variation begin and end abruptly (Fig. 4 of ref. 1) is not well supported by his data. Lengthy intervals of time may exist between those horizons exhibiting character deviance and those which precede and follow them. Nothing is known about the mode of evolution during these intervals. For example, Williamson's Figs 1 and 2 show that the first fauna which precedes fauna 8 and which contains *B. unicolor* (fauna 5b) lies 4.5 m below fauna 8. The next youngest fauna above 12c to contain *B. unicolor* (fauna 17) lies at least 30 m above the deposits containing highly deviant populations (fauna 17 actually lies much more than 30 m above fauna 12c, as an unspecified interval is shown between subsections 2 and 3 in Fig. 1). Similarly, in the next anomalous populations (faunas 27–29) no *B. unicolor* occur for about 12 m below or 1 m above deviant population 27.

Deviant population 88 has no stratigraphically close, nondeviant neighbours, and Fig. 4 is misleading in illustrating 'punctuated' shifts of morphology at this (the Guomde Formation) horizon; chronological gaps of unknown length lie at its upper and lower boundaries³ so that it is difficult to estimate evolutionary rates from a comparison of the Guomde populations with their neighbours. The faunas at the Guomde level do not "clearly document speciation within peripheral isolates" as Williamson claims.

Williamson's Fig. 1 shows the stratigraphic positions of his molluscan faunas and that they were collected from geographically separate areas with most of the faunas (including two morphologically deviant populations, faunas 27 and 29) stratigraphically located between

tuffaceous marker horizons. Williamson determined the stratigraphic placement of inter-tuff faunas by measuring the vertical distance from each faunal locality to the base of the nearest marker horizon, and assuming similar sedimentation rates in each area. However, sedimentation is known to be extremely variable in rate and character in lake deltas—a common palaeoenvironment in the north-east Turkana stratigraphic section^{4,5}. We therefore question the accuracy of the ordering faunas in this area. For example, faunas collected from 2 m above the same tuff horizon, in separate areas, could be of different ages if they occur in different interfingering lithological units. Widely separated faunas would be more susceptible to ordering error, and this reduced stratigraphic resolution might further reduce the capacity of this to document episodes of speciation.

Williamson believes that lake regression was the primary cause of speciation events in the Turkana Basin. However, the fossil evidence does not support the view that a major regression occurred at the Suregei Complex horizon. In area 202, the Suregei Complex contains the diatoms *Melosira granulata*, *Melosira agassizi* and *Cyclotella meneghiniana* (ref. 6 and J. Richardson, personal communication). In area 102 this complex comprises an open, deep-water (30 m) ostracod fauna, *Sclerocypris exserta* and *Gomphocythere* sp. Both of these are indicators of low to moderate alkalinity water⁶, and therefore suggest accumulation during a period of high (rather than low) lake level. The following model may be a more satisfactory explanation of the appearance of derived morphotypes at the Suregei Complex horizon.

Deep lake populations of many sexual and asexual molluscs are commonly more variable than are littoral populations^{7–9}. Organisms in deeper water tend to lose the dispersal capabilities found in shallow water species and normally exist as highly localized populations with restricted gene flow^{10–12} and increased morphological divergence and speciation. For example, in old deep lakes, cosmopolitan molluscan taxa are generally restricted to littoral environments, whereas more derived, descendant forms are more abundant in deeper water¹³.

Because most of the Turkana sediments represent marginal nearshore deposits, few of these ancient deep-water faunas have been observed. Williamson's 'speciation events' might represent incursions of deep water-derived morphotypes (perhaps even of different species) into shallow areas during brief periods of high lake level. Subsequent disappearance of such

forms might indicate a return to shallower conditions.

This model of the environmental conditions associated with the appearance of novel forms is quite different from that of Williamson. It suggests that although changes in lake level may coincide with morphological change they are not the causes of speciation. Instead it suggests that the occasional invasion of deep water populations into the East Turkana fossil records, means that it is impossible to use them to estimate the rate of morphological change and hence to distinguish between punctuated and gradual evolution.

ANDREW S. COHEN*
HILDE L. SCHWARTZ†

* Department of Geology,
Colorado College,
Colorado Springs, Colorado 80903, USA
† Earth Sciences Board,
University of California,
Santa Cruz, California 95064, USA

1. Williamson, P. G. *Nature* **293**, 437–443 (1981).
2. Williamson, P. G. *Nature* **296**, 611–612 (1982).
3. Findlater, I. C. in *Koobi Fora Research Project Monogr.* **1**, (eds Leakey, R. E. F. & Leakey, M.) (Clarendon, Oxford, 1978).
4. Findlater, I. C. thesis, Univ. London (1976).
5. Johnson, G. D. & Reynolds, R. in *Earliest Man and Environments in the Lake Rudolph Basin* (eds Coppens, Y., Howell, F. D., Isaac, G. & Leakey, R. E. F.) 115–122 (University of Chicago Press, 1976).
6. Cerling, T. E. thesis, Berkeley Univ. (1979).
7. Crowley, T. E., Pain, T. & Woodward, F. R. *Annals Mus. Afr. cent. Sér.* **8**.
8. Stankovic, S. *Int. Ass. theor. appl. Limnol.* **12**, 478–507 (1953).
9. Khozov, M. *Monogr. Biol.* **11** (1963).
10. Mayr, E. in *Evolution as a Process* (eds Huxley, J., Hardy, A. C. & Ford, E. B.).
11. Greenwood, P. H. *Bull. Br. Mus. nat. Hist. Zool. Suppl.* **6** (1974).
12. Brooks, J. L. *Q. Rev. Biol.* **25**, 30–60 131–176 (1950).
13. Brown, D. S. *Freshwater Snails of Africa and their Medical Importance* (Taylor & Francis, London, 1980).

THERE are two views of evolutionary pattern and process: either that most change occurs during relatively brief periods of speciation (the punctuation model) or that such changes occur gradually in established lineages (the gradualistic model). Williamson¹ concludes from his study of fossil freshwater molluscs in Lake Turkana that evolutionary patterns correspond to the punctuation model. A major problem in the taxonomy of freshwater molluscs is that of extensive intra- and interpopulation variability, which is presumably of ecophenotypic origin. Most morphological characters chosen by Williamson to describe shell shape are indeed very variable. In consequence, what he describes as the simultaneous origin of new lineages could be no more than the simultaneous origin of new environmentally induced variants. In addition, Williamson's reconstruction of events underlying speciation largely ignores much previous evidence.

The systematics of freshwater molluscs has recently greatly been clarified by the study of soft parts such as the reproductive and digestive tracts, the mantle edge and

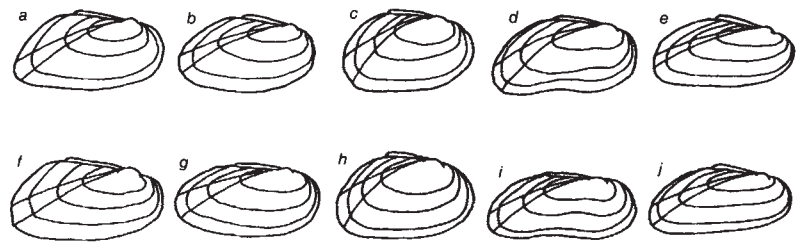


Fig. 1 An example of the wide range of phenotypic variability that can be present among species of the freshwater bivalve family Unionidae. In this case, variability expressed by *Elliptio complanata* (a–e) overlaps shapes typical of *Elliptio hopetonensis* (f), *Elliptio icterina* (g), *Elliptio congarea* (h), *Elliptio arcata* (i) and *Elliptio lanceolata* (j).

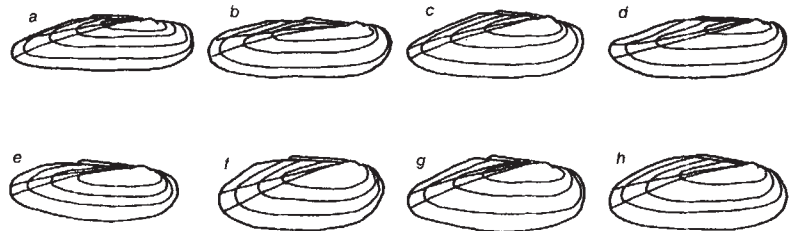


Fig. 2 Convergence on a lanceolate shape among unionid bivalves. a, *Elliptio shephardiana*; b, *Elliptio folliculata*; c, *Ligumia nasuta*; d, *Elliptio producta*; e, *Ligumia recta*; f, *Elliptio jayensis*; g, *Elliptio fisheriana*; h, *Elliptio lanceolata*.

the radular teeth^{2,3}. Enzymatic, chromosomal and shell microstructural analyses are also invaluable in determining taxonomic identity in difficult groups^{4–6}. The use of such taxonomic characters (which have high heritability) has had two effects on freshwater mollusc systematics; it has led to an appreciation of the degree of environmentally modified variability to which most freshwater molluscs are subject, and has led to the rejection of systems of classification based exclusively on shell characters such as shape, size, ornamentation or dentition. This poses a problem for fossil material because the taxonomically valuable characters are not preserved, and so identification must be based on characters with high levels of ecophenotypic variability.

Classification will then be prone to two errors; underestimation of the number of species in a fauna because of convergence of shell shape, and overestimation of species number because of extensive intraspecific variability (Fig. 1).

Convergence in shell shape and ornamentation is common among freshwater molluscs. For example, there is convergence on a particular shell sculpture and shape among gastropods of the families Hydrobiidae, Pomatiopsidae and Viviparidae^{2,7}; chromosomal and molecular genetic analyses demonstrate parallel evolution in shell shape in pleurocerid gastropods^{8,9}; and there is parallel evolution towards a lanceolate shape among species of the genus *Elliptio* (Fig. 2). Wetherby¹⁰ discussed intra-specific variability a century ago, and a multitude of factors such as sediment

type, water depth and composition, habitat type, wave exposure and currents, have now been shown to affect shell shape to a sufficient degree to cause species sampled from different habitats in the same drainage to exhibit a variety of non-overlapping phenotypes^{11–14}.

However, Williamson¹⁵ simply dismisses most of these problems of phenotypic variability. He devotes much attention to change in the *Bellamya unicolor* lineage, which is stated to correspond to patterns of change in the other lineage examined. The divergent forms of this species which arose during periods of lake regression are considered by Williamson to lie outside his conception of the narrow phenotypic range of *B. unicolor* and thus to constitute descendant species. However, it is clear in the context of the extensive information available on European and North American freshwater snails that the range of phenotypic variability expressed in these lineages is not at all narrow^{16–19}. In addition, the Lake Tanganyika viviparid *Neothauma tanganyicense*^{20,21} (which is closely related to *B. unicolor*) shows a range of variability within a single species in various habitats which completely overlaps that exhibited by the various phenotypes of fossil *B. unicolor* which Williamson considers to be distinct species (Fig. 3). Finally, Rohrbach²² examined several variants of *B. unicolor* itself using reproductive tract anatomy, and rejected shell shape as a valid means of classification for the group.

Williamson stresses that speciation in the Lake Turkana fossil fauna coincides

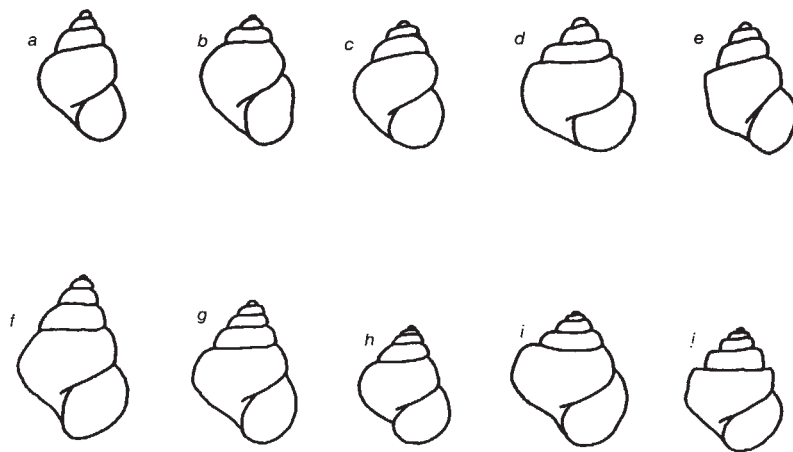


Fig. 3 A comparison of the range of variability in shape exhibited by *Neothauma tanganyicense* (f-i) and *B. unicolor* (a-e). Redrawn from Williamson¹ are shown the ancestral form of *B. unicolor* (a; faunas B and 3), several phenotypes that occur during periods of high variability at the Suregei level (b-d; faunas 8 and 11), and a 'daughter species' occurring just before a subsequent transgression (e; fauna 12). Below them are drawn a series of the closely related viviparid *N. tanganyicense* illustrated by Darteville and Schwet²¹ which represent ecophenotypes from several locations in Lake Tanganyika; collections at the Academy of Natural Sciences of Philadelphia (ANSP 131864, 140978) indicate both the existence of intermediates between the phenotypes illustrated, and that a series of phenotypes can occur at a single site. Similarly, examination of museum collections that acknowledge synonymies of *B. unicolor* reveals that several phenotypes such as e, c and a can occur at a single collection locality (ANSP 310346, 248839).

with lacustrine regressions, and invokes Mayr's model that speciation is most likely to occur in geographical isolates. However, Mayr²³ acknowledges that while geographical isolates might arise through contraction of a previously continuous species range into pockets, he points out that genetic divergence will occur only if the population involved suffers a great reduction in size. Williamson admits that population sizes in the lake were very large even during regression periods. It is therefore difficult to see why the claimed breakdown in homeostasis and developmental stability should have occurred.

It is also unlikely that all species should show genetic divergence from their ancestors at similar rates during lake regressions. The fauna consists of several very different lineages with different means of feeding and reproducing. Rates of divergence within different fossil mollusc lineages vary considerably, even among lineages which are undergoing adaptive radiation²⁴. Work on Unionidae⁴ indicates that such rates of divergence vary widely among clades.

These considerations together with the known ability of freshwater molluscs to respond phenotypically to environmental shifts, suggest that the simultaneous transformations in phenotype found in fossil lineages in Lake Turkana are better explained as a common response to an altered habitat than to simultaneous speciation events. These changes took 5,000–50,000 yr to accomplish. This could represent no more than the time necessary for a large body of water such as Lake Turkana to change from one habi-

tat type to another. Williamson's observation of morphological stasis over long periods of time is not incompatible with the snails' ability to show extensive ecophenotypic variation; if the populations sampled perceived a uniform habitat through time, overall similarity in shape is predicted.

More careful analyses of shell characters with high levels of heritability (such as conchiolin layer microstructure⁶) might help to establish the nature of shell shape changes in Lake Turkana. However, if the unit of macroevolutionary change is indeed the species as suggested by Williamson and by Gould, then it becomes essential to be able to recognize species in the fossil record. To disregard the information in the taxonomy of its modern members is greatly to reduce the value of Williamson's uniquely complete fossil assemblage.

PIETER W. KAT*
GEORGE M. DAVIS†

* Department of Earth and
Planetary Sciences,
Johns Hopkins University,
Baltimore, Maryland 21218, USA

† Department of Malacology,
Academy of Natural Sciences,
19th and the Parkway,
Philadelphia, Pennsylvania 19103, USA

1. Williamson, P. G. *Nature* **293**, 437–443 (1981).
2. Davis, G. M. *Acad. nat. Sci. Philadelphia Monogr.* **20**, (1979).
3. Kat, P. W. *Malacologia* **23**, 361–374 (1982).

4. Davis, G. M., Heard, W. H., Fuller, S. L. H. & Hesterman, C. *Biol. J. Linn. Soc.* **15**, 131–150 (1981).
5. Chambers, S. M. *Evolution* **36**, 113–120 (1982).
6. Kat, P. W. *Malacologia* **24**, (in the press).
7. Annandale, N. *Rec. geol. Surv. India* **50**, 209–243 (1919); *Proc. R. Soc. B96*, 60–76 (1942).
8. Chambers, S. M. *Malacologia* **17**, 157–162 (1980).
9. Dillon, R. T. & Davis, G. M. *Malacologia* **20**, 83–98 (1981).
10. Wetherby, A. G. *Am. J. Sci.* **23**, 203–212 (1882).
11. Ortmann, A. E. *Proc. Am. phil. Soc.* **54**, 269–312 (1920).
12. Baker, F. C. *Ecology* **9**, 271–283 (1928).
13. Brown, C. J. D., Clark, C. & Gleissner, B. *Am. Midl. Nat. Monogr.* **19**, 682–701 (1938).
14. Eagar, R. M. C. *Biol. Rev.* **53**, 169–210 (1978).
15. Williamson, P. G. *Nature* **296**, 611–612 (1982).
16. Pilbry, H. & Bequaert, J. *Bull. Am. Mus. nat. Hist.* **53**, (1927).
17. Haas, F. *Abh. Senckenb. naturforsch. Ges.* **431**, (1936).
18. Leloup, E. *Inst. R. Sci. nat. Belg.* **3(1)**, 1–146 (1950).
19. Brown, D. S. *Freshwater Snails of Africa and their Medical Importance* (Taylor & Francis, London, 1980).
20. Leloup, E. *Inst. R. Sci. nat. Belg.* **3(4)**, 1–168 (1954).
21. Darteville, E. & Schwet, J. *Mem. Inst. R. Coll. Belg.* **3(4)**, 1–128 (1948).
22. Rohrbach, F. *Arch. Molluskenk.* **69**, 177–218 (1937).
23. Mayr, E. *Populations, Species, and Evolution* (Harvard University Press, Cambridge, 1971).
24. Stanley, S. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 646–650 (1975); in *Patterns of Evolution* (ed. Hallam, A.) 209–249 (Elsevier, Amsterdam, 1977).

WILLIAMSON REPLIES—WITH regard to the points raised by Cohen and Schwartz, I am well aware of the importance of taphonomic considerations when making evolutionary interpretations. The Turkana Basin molluscs occur as three main facies, as discussed at length in my forthcoming monograph: life assemblages which have undergone no or little post-mortem transport; faunas emplaced as mass flows in prodeltaic settings; and 'death assemblages' consisting of clearly transported and in part abraded material. Only in the last category is a mixture of material from distant locations or of different ages a potential problem. Such faunas were avoided as far as possible in my analysis for precisely the reasons suggested by Cohen and Schwartz; although faunas from the Upper Member include several of this type they have little impact on my analysis as little of evolutionary significance occurs in the Upper Member¹. Most of my faunas are of the first two classes: postmortem transport is not a significant factor in this study (P. G. W., in preparation).

Even a superficial reading of my reply to Charlesworth and Lande² should indicate that my statement "the faunal units concerned seem to be undisturbed life assemblages" refers specifically to the transitional populations below the Suregei (faunas 7–11). With one possible exception (fauna 7), these faunas are indeed life assemblages.

Cohen and Schwartz are incorrect in suggesting that *Bellamya* is unknown in the section between faunas 12 and 17. As I make clear in my original article¹ and in my reply to Mayr², typical *Bellamya unicolor* is known from 2–3 m above fauna 12 (in fauna 13). This material was not included in the canonical variate analysis I presented; although well preserved, the material is heavily indurated and not amenable to the digitization technique used in this study to quantify

morphology. Fragmentary individuals of typical *B. unicolor* are also known faunas 15 and 16, and are readily distinguishable from the aberrant Suregei level *Bellamyia* species: the *Bellamyia* lineage is indeed documented over the interval faunas 12–17 (as are all other lineages used in this analysis).

The novel Suregei-level *Bellamyia* arises in only 5 m of section. Cohen and Schwartz do not regard this as particularly 'rapid'. The entire Turkana Sequence is around 500 m thick³ and represents some 3–4 Myr. Derivation of the Suregei level *Bellamyia* therefore occupies some 1–2% of the total stratigraphic range of the ancestral *B. unicolor* lineage in the Turkana sequence. In fact, typical *B. unicolor* is known in much older sections elsewhere in Africa (for example, it occurs at Lothagam, dated at >6 Myr)⁴. It is quite clear that the derivation of the Suregei level *Bellamyia* occurred extremely rapidly relative to the long duration of the ancestral *B. unicolor* lineage.

Although only one well-documented population of the distinctive Guomde level *Bellamyia* is known (from fauna 88, in the lowermost Guomde), additional fragmentary material is known from fauna 92, in the uppermost part of this formation. Of the other deviant taxa known from the Guomde, the endemic *Pseudobovaria* and *Melanoides* species are known from all Guomde faunas (87–92), and the endemic Mutelid is known from half of them (87, 90 and 92). I fail to see that Fig. 4 in my original article¹ is in any sense misleading. Events at both the Suregei and Guomde levels clearly reflect rapid, geographically localized, and profound morphological transformations of widespread, otherwise evolutionary conservative lineages—that is, punctuated equilibrium.

Although diatom faunas from the Suregei level in area 102 indicate a relatively open, non-alkaline lake as Cohen and Schwartz suggest, this is irrelevant to the conditions in the 5–10 m of section below the Suregei Complex where the evolutionary transformation from ancestral to descendant mollusc taxa is actually documented. However, it is noteworthy that neither Abell's recent oxygen isotope determinations on material from the Suregei Complex⁵, nor Brown and Cerling's recent suggestion³ of a major erosional disconformity immediately above the Suregei level in 102 would in fact support Cohen and Schwartz's palaeoenvironmental interpretations.

Cohen and Schwartz criticize my stratigraphic ordering of the mollusc faunas. In fact, over 70% of these faunas come from the following restricted collecting areas or units: area 202 (for faunas up to and including the Suregei level); areas 102 and 123–110 (faunas from the Lower Member (the Suregei level and above) and the Upper Member of the Koobi Fora Formation); the Guomde and the Galana Boi

Beds. Recent reappraisals of the stratigraphy of the Turkana sequence³ have not questioned the superpositional relationships of the sections in 202, 102 and 123–110, the Guomde and Galana Boi. Within areas 202, 102 and 123–110, faunas were deliberately collected from single, continuous, walkable sections. Although the details of the correlation of the Lower Member between areas 102 and 123–110 might be questioned, minor adjustments will hardly affect my conclusions as these Lower Member faunas simply record long periods of stasis or the abrupt appearance of new taxa without intermediate forms. Given the small number of sections from which the faunas I used were actually collected, and the unambivalent superpositional relationships of these sections, the relative stratigraphic positions of the faunas are not open to question. Finally, over a third of the faunas come from a single continuous section in a single area—102. This single run of faunas documents—albeit in incomplete form—all significant evolutionary events summarized in Fig. 4 of my article¹ up to the upper part of the Upper Member.

Cohen and Schwartz suggest that events at the Suregei and Guomde levels simply represent a facies-controlled invasion of a deep-water endemic fauna. There are several problems with their hypothesis:

(1) Their model requires an endemic lacustrine radiation in which all endemic taxa are totally excluded from littoral situations. But the literature they themselves cite indicates that no lacustrine radiation of this type—recent or fossil—is known.

(2) Cohen and Schwartz are confused as to the nature of lacustrine radiations. These radiations involve—by definition—a multiplication of taxa. Some fraction of the 'cosmopolitan' fauna in the basin gives rise to one or more coexistent endemic taxa. The Lower Member of the Koobi Fora Formation in the Turkana sequence clearly represents just such an adaptive radiation. But rather than radiation, the Suregei and Guomde level events document transformation of all taxa. Such a one-to-one transformation of ancestral to descendant taxa is a phenomenon quite distinct from the adaptive radiations to which Cohen and Schwartz refer—but might be expected within a peripheral isolate⁶.

(3) If the Suregei level novelties represent simple facies-controlled immigration of deep-water species derived from the long-ranging ancestral stocks, what are the intermediate morphologies immediately below the Suregei level supposed to represent?

Cohen and Schwartz accept the existence of long-term morphological stasis in the Turkana mollusc sequence, but appear unwilling to accept the relevance of this to the punctuated equilibrium–phyletic gradualism debate. But as authors too numerous to cite here have reiterated, the simple demonstration of

long-term morphological stasis—in and of itself—automatically calls the concept of traditional phyletic gradualism into question.

In reply to Kat and Davis: These authors suggest that the phenotypic transformations in the Turkana Basin molluscs are ecophenotypic modifications rather than speciation events. Certain groups of freshwater molluscs do indeed show considerable variation in shell morphology. However, the main point of my study is that there is *no* significant variation in any lineage for most of the 3–4 Myr represented by the Turkana section, with the exception of three brief episodes of morphological change (which occupy less than 3% of the total time represented by this sequence). If these various lineages are as ecophenotypically labile as Kat and Davis would have us believe, their stasis over 4 Myr could only be interpreted as the result of rigorous environmental stability throughout this immense span of time. But all previous palaeoenvironmental studies have stressed the profound environmental oscillations experienced at this and other East African sites during the later Cenozoic (for example 7–11). Clearly, 4 Myr of morphological stasis in 13 lineages cannot be explained away by simplistic appeals to "environmental stability".

The temporal stasis documented in the various lineages in the Turkana Basin sequence is mirrored by the striking geographical uniformity exhibited in a wide range of environments by their modern representatives (see, for example, Fig. 4). The ecophenotypic variation these exhibit at present is—like that portrayed in Fig. 1 of Kat and Davis' comments—trivial compared with the profound phenotypic reorganization documented at the Suregei, Guomde, and Lower Member levels in the Turkana Basin sequence.

Contrary to Kat and Davis' assertions, the work they cite on variation in shell shape with environmental conditions has emphatically not demonstrated non-overlapping phenotypes in any one species sampled from different habitats^{12–15}. Rohrbach's rejection of shell form as a valid means of classifying some *Bellamyia* species was based on a primitive morphometric analysis using the notoriously uninformative ratio of shell height/shell breadth¹⁶.

It may be that phenotypic variation in *Neothauma tanganyicense* encompasses the range of shell form seen in distinct *Bellamyia* species. Nevertheless, this is of peripheral relevance to the observation that the novel forms of *Bellamyia* from the Suregei Guomde and Lower Member levels in the Turkana sequence are as distinct from ancestral *B. unicolor* as are modern endemic *Bellamyia* spp. known to be 'good' biospecies.

The populations involved in the Suregei level speciation events are indeed too

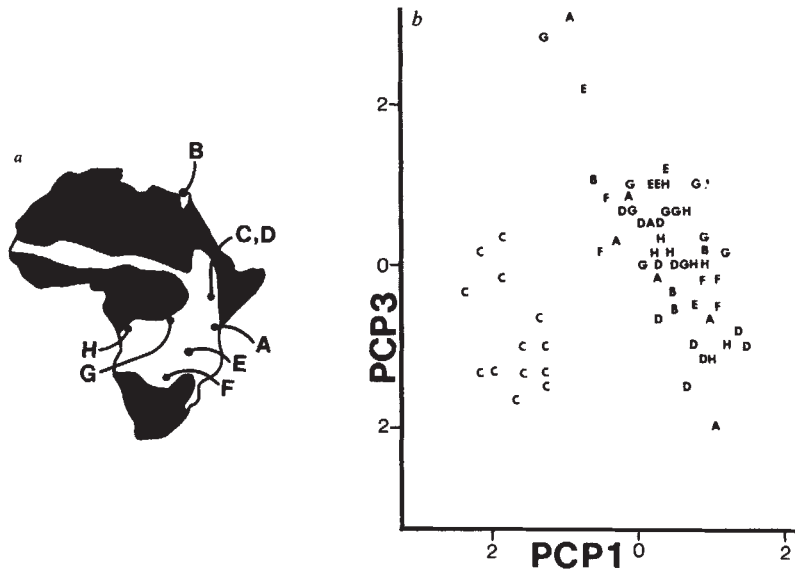


Fig. 4 *a*, Geographical location of the populations of *Bellamya* used in the Principal Components analysis reported in *b*. Light area on the map corresponds to the present distribution in Africa of the genus *Bellamya* (after Brown)¹⁸. Note that, although 16 species of *Bellamya* are currently recognized from Africa, all but two are endemics of limited distribution. The two widespread species—*B. unicolor* and *B. capillata*—may in fact be conspecific. The distribution indicated is therefore that of two, and perhaps only one, widespread African *Bellamya* species. A, Kat and Davis' population of *B. unicolor*, ANSP 310346, from south of Gama, Eastern Province of Tanzania. B, Kat and Davis' population of *B. unicolor*, var. *Biangulata* Kunst., ANSP 248839 from Alexandria, Egypt. C, Novel species of *Bellamya*, derived from *B. unicolor*, reported by Williamson¹ as population 12 from the Suregei tuff level in the Plio-Pleistocene Turkana Basin sequence. D, *B. unicolor* from the Lower Member of the Plio-Pleistocene Koobi Fora Formation (population 22B in ref. 1). E, *B. unicolor* from the Chambezi River, northern Zimbabwe (MCZ 109855). F, *B. unicolor* from Victoria Falls on the Zambezi, southern Zimbabwe (MCZ 56745). G, *B. unicolor* from Kindu, River Congo, northern Zaire (MCZ 77277). H, *B. unicolor* from the Congo River at Leopoldville, Zaire (MCZ 40900). *b*, Summary of Principal Components analysis of eight populations of modern and fossil *B. unicolor*, and the novel species derived from it at the Suregei Tuff level in the Turkana sequence. First Principal Component is plotted against the third. These components together explain 38% of the total variance. Letters A to G indicate the scores, on these components, of the various individuals from the eight populations whose geographic locations are indicated in *a*. The 19 morphometric parameters determined for each individual are: Raup's W (whorl expansion rate) for the last four whorls; Raup's T (translation rate) for these whorls; the relative height versus width of the generating curves of the last four whorls; the relative adapicality of the most abaxial point on the generating curve of these whorls; and the relative adapicality of Cox's X-point for the earliest three of these whorls. Note (1) that Kat and Davis' populations ANSP 310346 and ANSP 248839 (populations A and B in this analysis) group with typical *B. unicolor*; (2) that the novel form of *Bellamya* derived at the Suregei level of the Turkana sequence (population C) is quite distinct from *B. unicolor* (including Kat and Davis' populations); (3) the striking phenotypic uniformity of *B. unicolor* across its present extensive geographic range.

large for the founder effect to be a significant agent of change. The fact that large populations are involved in phenotypic transformations does not invalidate my contention that these are speciation events: strong selection in large populations—a likely concomitant of lake regression—can also disrupt homeostatic mechanisms¹⁷. Founder effect could not, in any case, trigger disruption in homeostasis in the ameiotic *Melanoides* at the Suregei level.

I did not state that all lineages in the Turkana sequence transform at the same rate during the Suregei episode. As I pointed out, the rate of transformation varies significantly. For example, below

the Suregei level, the peak in phenotypic variation (presumably a reflection of "disruption of homeostasis") in some lineages predates by 4 m and some 5–50,000 yr the comparable peak in other lineages.

The most serious charge levelled by Kat and Davis is their claim (Fig. 3) that the novel derived species of *Bellamya* from the Suregei level is a simple ecophenotypic variant of *B. unicolor* which commonly occurs in modern populations of this species, specifically in *B. unicolor* populations held in the collections at the Academy of Natural Sciences of Philadelphia (ANSP lots 310346 and 248839). What is the basis for this claim? Kat and Davis present no analysis of this material

and have never examined any of the Turkana Basin material. I have performed a morphometric analysis of these populations ANSP 310346 and 248839, which indicates quite unequivocally that they fall well within the rather narrow phenotypic range of *B. unicolor*; populations of *B. unicolor* from widely spaced locations in Africa, from the Turkana sequence and from Kat and Davis' populations, all group together. But the novel form of *Bellamya* from the Suregei level at Turkana is quite distinct from ancestral *B. unicolor* with which it forms a completely non-overlapping cluster. Contrary to Kat and Davis' assertions, no form comparable to this novel type occurs in population ANSP 310346 or 238839 (or in any other recent or fossil population of *Bellamya* outside the Suregei level faunas).

The highly variable populations immediately below the Suregei level, transitional between typical "ancestral" and derived forms, frequently include, in all studied lineages, both ancestral and derived morphologies, as well as truly intermediate forms and other variants (for example, Fig. 3 in ref. 1). Ecophenotypes are currently defined as "nongenetic modification(s) of the phenotype in response to an environmental condition" (ref. 19). Quite apart from the other considerations detailed above, the sympatric occurrence of "derived" and "ancestral" morphologies in these undisturbed transitional life assemblages clearly precludes any interpretation of the derived morphologies as simple ecophenotypes. In summary, available evidence strongly indicates that events at the Suregei and Guomde levels in the Turkana Basin sequence indeed document speciation events within peripatric isolates.

P. G. WILLIAMSON

Department of Geological Sciences,
Harvard University,
Cambridge, Massachusetts 02138, USA

- Williamson, P. G. *Nature* **293**, 437–443 (1981).
- Williamson, P. G. *Nature* **296**, 608 (1982).
- Cerling, T. E. & Brown, F. H. *Nature* **299**, 216–221 (1982).
- VanDamme, D. thesis, Rijksuniversiteit (1976).
- Abell, P. I. *Nature* **297**, 321–323 (1982).
- Williamson, P. G. *Nature* **295**, 140–142 (1982).
- Cerling, T. E. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **27**, 247–285 (1979).
- Findlater, I. C. in *Koobi Fora Research Project Monographs* 1 (eds Leakey, R. E. F. & Leakey, M.) (Clarendon, Oxford, 1978).
- Cerling, T. E., Hay, R. L. & O'Neil, J. R. *Nature* **267**, 137 (1977).
- Bonnefille, R. *Bull. Soc. Geol. Fr.* **21**, 331–342 (1979).
- Harris, J. M. in *Earliest Man and Environments in the Lake Rudolf Basin* (eds Coppens, Y. et al.) 293–301 (University of Chicago Press, 1976).
- Ortmann, A. E. *Proc. Am. phil. Soc.* **54**, 269–312 (1920).
- Baker, F. C. *Ecology* **9**, 271–283 (1928).
- Brown, C. J. D., Clark, C. & Gleissner, B. *Am. Midl. Nat. Monogr.* **19**, 682–701 (1938).
- Eagar, R. M. C. *Biol. Rev.* **53**, 169–210 (1978).
- Rohrbach, F. *Arch. Molluskenk.* **69**, 177–218 (1937).
- Levins, D. A. *Am. Nat.* **104**, 343–353 (1979).
- Brown, D. S., *Freshwater Snails of Africa and Their Medical Importance* (Taylor & Francis, London, 1980).
- Mayr, E. *Populations, Species, and Evolution* (Harvard University Press, 1971).

Mantle plumes and hotspots

HARRISON AND LINDH¹ have recently reported a new determination of the relative motion between the geomagnetic frame of reference and the hotspot frame of reference. While I do not disagree with their approach, I would like to point out that the hotspot frame of reference is not necessarily the same as the mantle plume frame of reference which is the frame of more fundamental geophysical

significance. As pointed out by Harrison and Lindh, a hotspot is the surface manifestation of a mantle plume. Recent work by Whitehead² has shown that when a rising plume passes through a region of shear, the plume can be deflected by as much as 60° from the vertical. For the Earth, where shear in the mantle may be occurring at a depth of 100–200 km, the geological features which are interpreted as the hotspot may be displaced 175–350 km from the location of the corresponding plume in the mantle. This displacement of 1.5–3.0° will not be important if the plate moves uniformly over the mantle plume. However, in their analysis, Harrison and Lindh dealt with data from the past 200 Myr. During this time, major reorganizations of plate motion have occurred so that significant errors can arise if the location of a mantle plume is assumed to be the same as its associated hotspot. For example, in the case of a plate whose direction of motion changed by 90°, the hotspot associated with a stationary mantle plume would exhibit an apparent displacement of 2.5–5.0° with respect to the geomagnetic frame of reference.

While these considerations do not alter the conclusions of Harrison and Lindh, they do emphasize the fact that there can be significant differences between the frame of reference defined by the location of hotspots and the frame of reference of the mantle plumes which give rise to them.

KENNETH L. VEROSUB

Department of Geology,
University of California,
Davis, California 95616, USA

1. Harrison, C. G. A. & Lindh, T. *Nature* **300**, 251–252 (1982).

2. Whitehead, J. A. *Geophys. J. R. astr. Soc.* **70**, 415–433 (1982).

HARRISON AND LINDH REPLY—The mechanism suggested by Verosub is certainly possible, and as he indicates would tend to degrade the information about mantle plume movements, especially during times of changing patterns of plate motion. However, it is possible that plumes may rise much faster than the simple Stokes' law equivalents used by Whitehead¹. This can happen if there is

significant heat transfer between the rising hot blob and the higher viscosity mantle material surrounding it, causing the mantle material close to the blob to become less viscous. In this case, the hot blob may rise at a rate much greater than that calculated from the simple Stokes' law equation. What is more, it is possible for the hot blob to ascend a large distance through the mantle while remaining essentially molten. These conclusions were reached recently by Ribe².

C. G. A. HARRISON
T. LINDH

Rosenstiel School of Marine and
Atmospheric Science,
University of Miami,
Miami, Florida 33149, USA

1. Whitehead, J. A. *Geophys. J. R. astr. Soc.* **70**, 415–433 (1982).

2. Ribe, N. *Pap. at Conf. Seamounts*, Lamont-Doherty Geological Observatory (1982).

Two cubic phases in monoolein–water system

LONGLEY AND MCINTOSH¹ have given convincing evidence for the existence of a cubic phase with a primitive lattice in the monoolein–water system, the structure proposed consists of a lipid bilayer forming a tetrahedral periodic minimal

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 500 words and the briefest of replies, to the effect that a point is taken, should be considered.

samples prepared over the whole composition range appeared homogeneous, the other X-ray data were indexed according to the same lattice. The demonstration of the existence of a primitive lattice¹ has initiated a more thorough analysis of the X-ray data, and the new indexing of our earlier published X-ray data² is given in Table 1. The X-ray data of samples with 34.93 and 39.87% (wt/wt) of water are in perfect agreement with the data repor-

Table 1 New indexing of the X-ray diffraction lines recorded at 22 °C from cubic monoolein–water phases of different compositions.

Body-centred lattice ²						Primitive lattice ⁴					
25.05% (wt/wt) water			29.79% (wt/wt) water			34.93% (wt/wt) water			39.87% (wt/wt) water		
<i>d</i>	<i>a</i> _{calc}		<i>d</i>	<i>a</i> _{calc}		<i>d</i>	<i>a</i> _{calc}		<i>d</i>	<i>a</i> _{calc}	
1(Å)	(<i>hkl</i>)	(Å)	1(Å)	(<i>hkl</i>)	(Å)	1(Å)	(<i>hkl</i>)	(Å)	1(Å)	(<i>hkl</i>)	(Å)
48.9	211	119.8	54.5	211	133.5	59.8	110	84.6	64.2	110	90.8
42.2	220	119.4	47.0	220	132.9	49.3	111	85.4	52.2	111	90.4
32.8	321	122.7	36.1	321	135.1	43.4	200	86.8	46.0	200	92.0
30.8	400	123.2	33.9	400	135.6	35.2	211	86.2	32.6	220	92.2
27.3	420	122.1	30.1	420	134.6	30.3	220	85.7	30.7	221	92.1
26.1	332	122.4	29.1	322	136.5	28.7	221	86.1	29.3	310	92.7
25.0	422	122.5				27.1	310	85.7			
24.1	431	122.9									

surface which separates two water channel networks. According to earlier work² the cubic monoolein–water phase forms a body-centred lattice. I now show that there are, in fact, two cubic phases in this binary system, one with a body-centred lattice at low water content and another with a primitive lattice at high water content.

Our structure analysis in the monoolein–water system was based on NMR diffusion measurements and X-ray diffraction changes at the transition from a lamellar liquid–crystalline phase to a cubic phase obtained by heating at a low water content². The X-ray diffraction lines observed there were in agreement with a body-centred lattice, and as all cubic

ted by Longley and McIntosh¹. Note that the spacings and intensities at 20.05 and 29.79% (wt/wt) of water are in agreement with the data of the body-centred cubic phase described by Luzzati *et al.*³. The transition between the body-centred and primitive lattice at 22 °C takes place at a composition in the range 30–35% (wt/wt) of water.

K. LARSSON

Food Technology Chemical Centre,
Box 740, S-220 07 Lund 7,
Sweden

1. Longley, W. & McIntosh, J. *Nature* **303**, 612–614 (1983).

2. Lindblom, G., Larsson, K., Johansson, L., Fontell, K. & Forsén, S. *J. Am. chem. Soc.* **101**, 5465–5470 (1979).

3. Luzzati, V., Tardieu, A., Gulik-Krzywicki, T., Rivas, E. & Reiss-Husson, F. *Nature* **220**, 485 (1968).

Mineral devotion

Peter J. Smith

Geophysics in the Affairs of Man: A Personalized History of Exploration Geophysics and Its Allied Sciences of Seismology and Oceanography.

By Charles C. Bates, Thomas F. Gaskell and Robert B. Rice.

Pergamon: 1982. Pp.492. Hbk £33, \$60; pbk £12.50; \$25.

ONE of the more curious episodes in the history of geophysics was the way in which the Society of Jesus (the Jesuits) unwittingly aided the front-line defeat of the Germans during the First World War. When in 1910 the United States Congress rejected a proposal to establish a Bureau of Seismology at a cost of a mere \$20,000 a year, the Society offered its communities around the world as bases for a standardized seismic network with the aim, at least in part, of demonstrating that the Church was not the enemy of progress and enlightenment it had often been made out to be. It then proceeded to develop the scheme with such enthusiasm that within a year seismic observatories were being operated or built at no fewer than sixteen Jesuit colleges throughout North America. By October 1914, however, the experience thus gained in detecting earthquakes was being put to use on the Western Front to locate German heavy artillery, an irony compounded by the fact that the Jesuit network had been set up in the first place with low-cost seismographs made in Göttingen.

Bates and his colleagues recount this story not simply as an early example of the interaction between geophysics and society but because the incident led directly to the development of exploration geophysics, a phenomenon that was to change the world no less profoundly than the War itself. Once hostilities were over, four Americans who had been involved in designing the equipment for artillery detection concluded that reflected sound waves also had potential for detecting underground petroleum-bearing structures, worked out a possible method, built the necessary equipment, and formed the Geological Engineering Company. The company itself soon failed, passing into other hands during a temporary financial crisis; but there was no stopping the technology. The first seismic reflections were obtained on 12 April 1919 in a Maryland quarry and by July depths to subsurface discontinuities were being

determined in Oklahoma.

To say that subsequent results were impressive would be an understatement. By 1924 an oil-bearing salt dome in Texas had for the first time been discovered by seismic means alone. In that year too, another salt dome was detecting using only newly-developed gravity anomaly techniques. Thereafter, geophysical methods began to reveal sources of oil at an increased rate and with spectacularly improved economy. Eleven oil-bearing salt domes were discovered in 1928–1929 at a cost of \$50,000 each, whereas in the same region a discovery using geology alone had

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

The results of early oil exploration — suburban Los Angeles in 1895.

cost \$20 million only four years earlier. Proven resources of oil in the United States, which had increased roughly linearly over the first quarter of the century, continued to increase linearly from the 1920s but at three times the rate. Not all geophysical exploration was being carried out in America, of course; but the emphasis was there, with the result that by 1940 the United States was producing about two-thirds of the world's petroleum at a cost of \$1 a barrel.

It is one of the ironies of life that some of the most significant developments in society are not necessarily the most intellectually inspiring, except perhaps to those intimately involved; and although the story of the rise of the exploration industry

between the wars is told ably enough in this book, I cannot pretend I found it anything but mildly tedious. It is less the history of ideas than that of a ruthless devotion to a single idea. It may be in part the history of technological development, but it is no less the history of business and all that means in terms of rivalries, changing allegiances and the pursuit of profit by unheard-of characters with whom it is difficult to have much fellow feeling.

With the coming of the Second World War, however, the character of geophysical activity changed, hardly surprisingly. This was the war that some historians have chosen to call the "war of the physicists" but which Bates and his colleagues convincingly demonstrate could with no less justice be termed the war of the geophysicists. Exploration continued, but on a much reduced scale, whilst some of the geophysicists previously involved applied their skills to the design and testing of equipment of more immediate military value and others took to the location of artillery much as had been done in the earlier war. But the luckier ones had the opportunity to apply their experience of acoustics, magnetics and gravity anomalies in more novel ways, to mine and submarine detection. As a peace-loving citizen I'm sorry to have to say it, but the extended account of the ingenuity with which Allied geophysicists rapidly neutralized each new advance in German mine technology makes the most gripping reading in the whole book.

The most profound effect of the War on geophysics was yet to come, however, and in ways almost too numerous to mention. For one thing, some of the techniques developed rapidly for military purposes found almost immediate post-

War civilian application, not the least being aeromagnetic surveying with its ability to generate cheap reconnaissance data. Then many of the people who had for one reason or another found themselves working in wartime geophysics decided to stay with this hitherto unfashionable subject; Hess, Reville, Ewing, Munk, Vacquier, Bullard and a host of others were to become leaders of the profession, thereby raising its prestige as they speeded its progress. Next, the War had convinced governments of the need for even more, and especially meteorological, data, which had the side effect of opening up vast new regions (for example, the Arctic) to geophysical exploration. And with that came money, masses of money — so much of it, indeed, that by

A Shell photograph

1949 the National Academy of Sciences was complaining that the cash being thrown at oceanography was excessive. Finally, the money spawned new agencies, new companies and new university departments by the score, turning the United States in particular into a vast geophysical research and exploration machine the like of which had never been seen before.

The consequences were a wonder to behold. Exploration geophysics saw technical advance after technical advance, not least in the field of data processing. The more academic version of the subject went through a conceptual revolution with an impact greater than anything that had happened since catastrophism was overthrown nearly two centuries before. Geophysics entered the realm of Big Science with the International Geophysical Year, huge seismic arrays, deep-sea drilling and the monitoring of the Earth from space. And not least, geophysicists found themselves in politics — generally, for big projects require big pipers who tend to call big tunes, and specifically, because science-derived political problems such as nuclear control demand scientific solutions such as explosion-earthquake discrimination.

All this and more (at least insofar as it relates to North America and Britain) is documented in meticulous detail with copious footnotes, charts and tables, which makes for some heavy reading at times. Those of a particular political

persuasion will also find it irritating reading, for the viewpoint is hardly neutral. Bates and his colleagues make it quite clear that their aim is to document "the all-important role played by the innovator-capitalist operating in a free-enterprise system that has led to a major technology and service industry of great importance to the world's economic well-being". In short, the hard information comes in a politico-philosophical wrapper.

The result is as interesting to students of human nature as it is to geophysicists. The prevailing tone is one of old-fashioned technological optimism until, that is, we get to the 1970s and the rise of the "environmentalists". At this point Bates, Gaskell and Rice begin to come over as three rather nice old gents who have devoted their lives to the economic well-being of society (and have done pretty well out of it, as they would admit) but are now deeply puzzled, and perhaps even hurt, to find their basic values questioned by an unconventional enemy from within. They simply cannot comprehend the philosophical shift that has taken place; but then none of us yet knows just how deep or permanent it really is, or even whether it is a matter of genuine choice or historical imperative. □

Peter J. Smith is Reader in the Department of Earth Sciences at the Open University, and editor of Open Earth.

Talking physics

Ilya Prigogine

Physics as Natural Philosophy: Essays in Honor of Laszlo Tisza on His Seventy-fifth Birthday.

Edited by Abner Shimony and Herman Feshbach.

MIT Press: 1983. Pp.448.
£45, \$67.

DOCTORS Shimony and Feshbach are to be congratulated on having put together such a stimulating book to mark Professor Laszlo Tisza's seventy-fifth birthday. This gives me the opportunity to add my own congratulations. I remember my stay as Visiting Professor at Harvard in 1955, during which I had many discussions with him. We both were looking for a new path to take us beyond classical thermodynamics: even if our choice of routes was somewhat different, those discussions have been a source of inspiration for me.

The book contains five sections, each covering different aspects of Tisza's interests: the foundations of probability and thermodynamics; condensed matter physics; quantum mechanics and relativity; biological systems; and the history and philosophy of science. As a result the book provides a good introduction to some

of the salient problems which have preoccupied scientists throughout the twentieth century but I have two criticisms.

Only one article by Abner Shimony deals directly with Tisza's own work. The book includes a list of his publications, and some of these have been of lasting influence — for instance the papers on quantum hydrodynamics and on critical fluctuations. So it is a pity they are not considered in detail since this was an ideal opportunity to present this work in a historical perspective. I especially admire his theory of critical fluctuations, which is forerunner of the so-called "renormalization group theory" and whose basic approach is incorporated in the recent work on non-equilibrium fluctuations of my colleagues Nocolis, Malek Mansour and others.

My second criticism is to do with the uneven level of the contributions in the book. For example, at the end of Suzan J. Benofy and Paul M. Quay's long article ("Fourier's Inequality as a Principle of Thermodynamics with Applications to Magnetothermoelectric Effects"), one reads with surprise:

The fact that contemporary engineering seems to have found no problems with its heat-flow equations, though these are not restricted to the steady state, confirms the reasonableness of taking Fourier's inequality as a thermodynamic principle that can be sustained in its own right, unless, perhaps, someday someone succeeds in subsuming it under the Second Principle.

Most physicists would, I think, consider Fourier's inequality as nothing but an expression of the Second Principle. This is the point of view taken, among others, by de Groot and Mazur in their well-known book on non-equilibrium thermodynamics. That book also contains an excellent chapter on magneto-thermoelectric effects; if the authors disagree with such authorities they should make explicit their reasons.

In the same vein, the article by Erwin Oppenheim, "Thermodynamics and Fluctuations in Nonequilibrium Systems", presents as a kind of novelty the fact that non-equilibrium fluctuation cannot be expressed in terms of the usual thermodynamic fluctuation theory. This however has been well known for more than ten years now. (Strangely enough, no references are given except to the work of Oppenheim in course of publication!)

Another point of contention lies in the interesting paper by David Layzer, "Quantum Mechanics, Thermodynamics, and the Strong Cosmological Principle". Here Layzer argues that the strong cosmological principle leads to an interpretation of the second law of thermodynamics that incorporates the "hierarchy" of order-generating processes. Even if one could agree with Layzer that our description of the physical Universe has to be in terms of distribution functions, rather than individual trajectories or wave functions, one is still very far from a theory of entropy increase or order generation. It is true that every theory of irreversibility has to be compatible with cosmology but it is doubtful if any cosmological principle by itself can lead to a microscopic theory of irreversible processes. One should not forget that in our expanding Universe we have both some processes which we can describe as reversible — at least at a very high accuracy — and other processes which are strongly irreversible. Of course, as Roger Penrose has suggested, it may be that we have to add some quantum gravity terms to explain irreversibility, but that is still not clear. I believe that irreversibility is a manifestation of the fact that we have to go from the initial representation of classical or quantum mechanics which is expressed in terms of canonical or unitary transformations, to a new representation in which the second law is incorporated and so to a new kind of transformation theory dealing with a class of non-unitary transformation. The same applies to general relativity.

Fortunately, there are many excellent articles in the book. I particularly liked those by J. M. Luttinger and Elliott H. Lieb. Silvan S. Schweber's article is excellent and emphasizes the enduring truth that novel ideas often appear when points of view originating in one field of knowledge diffuse to another. □

Ilya Prigogine is Regental Professor at the University of Texas, Austin, and a Professor at the Université Libre de Bruxelles.

Broadening the perspective

Geoff Bailey

European Economic Prehistory: A New Approach.

By Robin Dennell.

Academic: 1983. Pp.214. £15, \$26.

THEMATIC approaches to prehistoric archaeology are no longer a novelty. Analysis of general issues rather than reconstruction of particular archaeological sites, areas and periods, and an emphasis on generalization and explanation rather than the relentless cataloguing of human cultural diversity for its own sake, are now well-established ideals of modern archaeological research. Few specialist reports fail to mention the ecological, social or cognitive implications of their subject matter, even if such comment is often little more than lip service to current intellectual fashion.

Yet general accounts of prehistory have tended to convey a quite different impression. The ever-growing mass of archaeological data has resulted in syntheses focused on ever smaller areas and periods, or in global treatments in which the continuum of prehistory, for want of any better organizing principle, is divided into social and technological stages of development inherited from nineteenth-century preconceptions. Palaeolithic Savages and Neolithic Barbarians, to say nothing of the ghostly ancestors of Civilization, still stalk the pages of the modern literature, lurking behind the lengthening lists of sites and radiocarbon dates, and stones and bones, and pots and artefacts of every sort.

Robin Dennell's *European Economic Prehistory* provides a long overdue break with nineteenth-century conventions in its attempt to provide an overview of central themes in European prehistory from the earliest human occupation over 500,000 years ago up to the spread of farming during the Neolithic period. It is neither wholly about Europe, since most issues during this time span cannot adequately be understood except by reference to a broader geographical context, nor is it solely about prehistoric economies. As Dennell emphasizes, economic prehistory should not be merely a description of past diets, nor should it be the expression of a narrow materialist philosophy. Rather it provides a unifying perspective on a whole range of issues in human physical and cultural evolution.

The earliest occupation and subsequent expansion of human settlement, the appearance of modern social and cognitive patterns, the responses of late Pleistocene populations to the extreme environments of the Last Glacial and the subsequent deglaciation, and the adoption of novel

food resources and the spread of farming economies — these are the highlights of Dennell's approach. Theoretical concepts are combined with a clear statement of the practical difficulties and controversies involved in the interpretation of primary archaeological and palaeontological data; compendious detail is avoided without courting superficiality or eccentricity. The conventional subdivisions of the Palaeolithic period are critically examined and found wanting as useful markers of behavioural change, while the attack on the Neolithic revolution as a useful concept in European prehistory is short, sharp and lethal.

This is very much a personal view — the ideas incorporated within the book will stimulate the specialist as well as the general reader, who will find in it a short but informed account of changing perspectives on our prehistoric origins. It should do much to enlighten those for whom the long stretches of the Pleistocene period seem like a dismal dark ages before the rise of civilization. □

Geoff Bailey is a Lecturer in the Department of Archaeology, University of Cambridge.

Hamburger heaven

Robert Ubell

The Social Transformation of American Medicine.

By Paul Starr.

Basic Books: 1982/Harper & Row: 1983.

Pp.514. \$24.95, £21.50.

MOST readers of *Nature* will be aware of the unusual organization of health care in the United States — a medical practice largely in the hands of private ("fee-for-service") doctors. Still, it is worth repeating for American readers that medical care in most other advanced capitalist nations is delivered by physicians serving under national health insurance schemes. Only in America does the siren of an ambulance rushing to save your life also signal the need to pull out your cheque-book or private insurance card.

The struggle for national health insurance in the United States and its relentless suppression is one of the principal motifs in this remarkable book. Paul Starr, a Harvard sociologist, reveals the forces at work in the creation of American medicine; acting like a geologist studying plate tectonics, he uncovers broad changes, shifts and displacements that helped mould the American medical landscape. By charting social and economic struggles, Starr draws a kinetic map showing how the private practitioner came to occupy the peak of power.

Starr frames his analysis by going back to the colonial period, when health care

was largely the province of women who routinely treated their families at home. As midwives, women assumed primary responsibility for childbirth. Prominent as "lay practitioners", women also frequently entered the medical profession as the equals of men. By 1893–1894, women accounted for ten per cent or so of students at co-educational medical institutions, in addition to taking all of the places at women's medical colleges. In some cities — Boston, for one — nearly twenty per cent of those practising were women. Strikingly, by the turn of the century, male doctors had taken over almost completely. Just as women cooks are usurped by male chefs when status and economic gain are at stake, so medicine became a male preserve.

Toppling any misconception that doctors have always held a dominant position in America, Starr shows that a medical career in the middle of the nineteenth century guaranteed neither prestige nor security. In 1832 on learning that his son was about to become a doctor, one father wailed:

If I had known this, I certainly would not have sent you to college . . . it is a profession for which I have the utmost contempt. There is no science in it. There is no honour to be achieved in it; no reputation to be made.

Throughout most of the century, American populism had opposed whatever privilege and social distinction doctors claimed. And because academic medicine had not yet secured any decisive victories, many Americans held that it was wiser to treat illness at home rather than turn to physicians who knew as much as they did. While resistance to professional medicine has often been portrayed as hostility to science, Starr contends that popular scepticism was hardly unreasonable, given what we know about the effectiveness of early nineteenth-century therapy. But as medicine moved into the new scientific orbit, it could finally be defended as therapeutically valuable. The "democratic" ideal — that clinical knowledge should be easily accessible and universally practised — shrank in the face of the growing complexity and efficacy of medicine. Americans, by the 1920s, had changed their minds. They no longer believed they could heal themselves without professional intervention.

By the end of the first part of this intricately argued book, Starr has pursued his story through to the Great Depression. By then, American physicians had secured their unprecedented authority — and more. "Acknowledged skills and cultural authority", Starr remarks, "are to the professional classes what land and capital are to the propertied. They are the means of securing income and power."

How did they do it? Other sectors of American society had fallen under the sway of the corporation or belonged to the state. For Starr, the answer, while not simple, is the hinge on which the door swings open

to our understanding of the social transformation of medicine in America.

The medical profession's success was won by achieving a broad consensus among practitioners. Divided in the previous century, they rose to strength through the single-mindedness of the American Medical Association, among other professional vehicles. Taking advantage of their newly gained acclaim, they used the law to suit their ends. And the law gave them nearly complete authority over every aspect of medical delivery.

According to Starr, however, it was neither the advance of science nor the doctors' monopoly over medical practice alone that secured the profession's control:

Science may improve efficacy and productivity of a profession without making it rich and revered; knowledge must be transferred to authority, and authority into market power, before the gains from scientific advance can be privately appropriated by a profession If the medical profession were merely a monopolistic guild, its position would be much less secure than it is With widespread support, which they received because of complex changes overtaking the entire society, physicians were able to see social interests defined so as to conform to their own [my italics]. This was the essence of their achievement.

Paradoxically, the success of the medical profession may be its eventual undoing. Until quite recently it was almost solely the doctors themselves who determined the relationship of patients to hospitals, provision of drugs and payments for

medical services. But while they have controlled most aspects of the marketplace, they have failed conspicuously to control costs. Doctors made health care lucrative, and large-scale enterprises have moved to take advantage. So while the medical profession has historically blocked government entry into health care by the front door, corporations have been moving in by the back. For example, according to the president of Humana (one of *Fortune's* 500 service companies) his organization aims at providing a uniform, reliable product, just like "MacDonald's hamburger coast to coast". By 1980, Humana, which had started with a few nursing homes, owned 92 hospitals and had \$1.4 billion in revenues.

The contemporary message of Starr's book is both clear and stark: the advent of corporations into medicine will undoubtedly further aggravate inequality in access to health care. "Profit-making enterprises are not interested in treating those who cannot pay . . . the two-class system in medical care is likely to become only more conspicuous."

This is a radical book, written in a patrician style: scholarly and balanced, yet ultimately subversive. It is a major accomplishment. If Starr doesn't take it on as his next task, someone should surely use this book as a model to write a similarly perceptive history of science in America. □

Robert Ubell, until recently the American publisher of Nature, is head of a science publishing house in New York.

Inside yeast cells

J.M. Mitchison & P.A. Fantes

The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression.

Edited by Jeffrey N. Strathern *et al.*
Cold Spring Harbor Laboratory: 1983.
Pp.680. \$75 (US), \$90 (elsewhere)

THE Cold Spring Harbor Laboratory has done as well with this second part of its two-volume work on yeast as it did with the first on life cycle and inheritance (for review see *Nature* 92, 299; 1982). It is a distinguished collection of review articles which centre round metabolism and which show how much progress has been made since the last major review of the field in 1971 (*The Yeasts*, edited by A.H. Rose and J.S. Harrison, and published by Academic). Inevitably, some of the articles are already dated, especially in subjects like recombinant DNA technology. But in the slower-moving areas the reviews should be definitive for a respectable period.

A common theme is the powerful link between genetics, biochemistry and cell biology which is a characteristic of modern work on budding yeast. This is well

brought out in the elegant article on secretion and surface assembly by Schekman and Novick. Another recurring problem is the complexity of many aspects of regulation and transport; anyone who has the temerity to think that regulation is simple should read the review by Oshima on the phosphatases.

The longest and perhaps the most impressive article is by Jones and Fink on amino-acid and nucleotide biosynthesis. There are clear accounts of pathway-specific and cross-pathway controls, and the tables and charts alone are (almost) worth the price of the book. Even so, at the end of more than a hundred pages we are told that "In no case do we know what the molecular basis of regulation is".

Those who work with budding yeast will find this book of great use. But it will be even more valuable to those who work with other fungi and with higher cells. More is known about budding yeast than about any other lower eukaryote and the natural question for those like ourselves who work with other yeasts is "what happens in budding yeast?". This book and its companion volume will tell us. □

J.M. Mitchison is Professor and P.A. Fantes a Lecturer in the Department of Zoology, University of Edinburgh.

Understanding NMR techniques

K.A. McLauchlan

Nuclear Magnetic Resonance Spectroscopy: A Physicochemical View.

By Robin K. Harris.

Pitman: 1983. Pp.250. £17.50, \$34.95.

TO WRITE any book on NMR spectroscopy is to expose oneself to criticism — the subject is so wide, both in itself and in its applications, that tough decisions have to be made on the approach to be taken. The problems lie in identifying the intended readership and then in deciding the depth into which one should go. Such difficulties are compounded by the continual emergence of fresh concepts and experimental techniques.

Professor Harris has made a deliberate and brave choice to bias his book away from the straightforward, normal applications of NMR and towards introduction of the physical concepts involved in even the most modern experiments. As one would expect, the writing is particularly clear and the contents unusually well thought out.

The book is subtitled *A Physicochemical View*, although for the most part a physical view is given with chemistry being used for illustration. Indeed physicochemical triumphs — including precise molecular geometry determination in solution and the nuclear polarization experiments which allow unique insight into free radical reaction mechanisms — are omitted deliberately. I was sorry too that a decision was made not to include density matrices, even though some of the subjects covered, notably exchange and two-dimensional spectroscopy, require them if the reader is to advance much beyond the level given here.

The book is somewhat didactic in style, with the origins of many equations being indicated only broadly; a dedicated magnetic resonance spectroscopist would need to consult more advanced monographs. Nor is it suitable as an introduction to NMR which an undergraduate could read to obtain an overview of the subject, its importance in chemistry and the analysis of simple spectra; it is too detailed for beginners. However for someone who has already progressed to that stage a firm basis for understanding in depth is provided and horizons are extended. Difficult ideas are met head on and illuminated. For those just getting to grips with nuclear magnetic resonance spectroscopy, the book will be most useful. □

K.A. McLauchlan is a Lecturer in Physical Chemistry at the University of Oxford, and a Fellow of Hertford College.

Astronomy

- HARTMANN, W.K. *Moons and Planets*, 2nd Edn. Pp.509. ISBN 0-534-00719-8. (Wadsworth: 1983.) Np.
- MURRAY, C.A. *Vectorial Astronomy*. Pp.350. ISBN 0-85274-372-6. (Adam Hilger: 1983.) £25.
- ROY, A.E. and CLARKE, D. *Astronomy: Principles and Practice*, 2nd Edn. Pp.342. Hbk ISBN 0-85274-463-3; pbk ISBN 0-85274-464-1. (Adam Hilger: 1982.) Hbk £20; pbk £9.95.
- ROY, A.E. and CLARKE, D. *Astronomy: Structure of the Universe*, 2nd Edn. Pp.296. Hbk ISBN 0-85274-465-X; pbk ISBN 0-85274-466-8. (Adam Hilger: 1982.) Hbk £20; pbk £9.95.
- SHUTER, W.L.H. (ed.). *Kinematics, Dynamics and Structure of the Milky Way*. Proceedings of a Workshop held May 1982, Canada. Pp.392. ISBN 90-277-1540-8. (Reidel: 1983.) Np.

Physics

- BATES, D. and BEDERSON, B. (eds). *Advances in Atomic and Molecular Physics*, Vol.18. ISBN 0-12-003818-8. (Academic: 1982.) \$64.
- BRANSDEN, B.H. *Atomic Collision Theory*, 2nd Edn. Pp.511. ISBN 0-8053-1181-5. (Benjamin/Cummings: 1983.) £22.45.
- BURKE, P.G., EISSNER, W.B., HUMMER, D.G. and PERCIVAL, I.C. (eds). *Atoms in Astrophysics*. Pp.356. ISBN 0-306-41097-4. (Plenum: 1983.) \$49.50.
- DEVREESE, J.T. and BROSENS, F. (eds). *Electron Correlations in Solids, Molecules, and Atoms*. NATO Advanced Study Institute, Vol.81. Pp.435. ISBN 0-306-41027-3. (Plenum: 1983.) \$55.
- DEVREESE, J.T., VAN DOREN, V.E. and VAN CAMP, P.E. (eds). *Ab Initio Calculation of Phonon Spectra*. Pp.303. ISBN 0-306-41119-9. (Plenum: 1983.) \$42.50.
- FLAMM, D. and SCHÖBERL, F. *Introduction to the Quark Model of Elementary Particles* (in two volumes), Vol.1: *Quantum Numbers, Gauge Theories and Hadron Spectroscopy*. Pp.372. ISBN 0-677-16270-7. (Gordon & Breach: 1982.) \$73.50.
- GOLDIN, E. *Waves and Photons: An Introduction to Quantum Optics*. Pp.211. ISBN 0-471-08592-8. (Wiley: 1982.) £19.50, \$31.
- GREINER, W. (ed.). *Quantum Electrodynamics of Strong Fields*. NATO Advanced Study Institutes Series, Vol.80. Pp.902. ISBN 0-306-41010-9. (Plenum: 1983.) \$110.
- HAWKES, P.W. (ed.). *Advances in Electronics and Electron Physics*, Vol.59. Pp.306. ISBN 0-12-014659-2. (Academic: 1982.) \$47.50.
- KRANE, K.S. *Modern Physics*. Pp.512. ISBN 0-471-07963-4. (Wiley: 1983.) £19.40, \$31.
- LARSEN, D.C. (ed.). *Thermal Conductivity*, Vol.16. Pp.623. ISBN 0-306-41032-X. (Plenum: 1983.) \$79.50.
- LERNER, R.G. and TRIGG, G.L. (eds). *Concise Encyclopedia of Solid State Physics*. Pp.311. Hbk ISBN 0-201-14204-X; pbk ISBN 0-201-14205-8. (Addison-Wesley: 1983.) Hbk £29.60; pbk £17.20.
- NOVIKOV, S.P. (ed.). *Mathematical Physics Reviews*, Vol.3 (1983). Pp.311. ISBN 3-7186-0107-9. (Harwood Academic Publishers, Switzerland: 1982.) Np.
- PAGELS, H.R. *The Cosmic Code: Quantum Physics as the Language of Nature*. Pp.370. ISBN 0-7181-2217-8. (Michael Joseph Ltd: 1983.) £10.95.
- TSUBOI, C. *Gravity*. Pp.254. Hbk ISBN 0-04-551072-5; pbk ISBN 0-04-551073-3. (George Allen & Unwin: 1983.) Hbk £20; pbk £9.95.
- YOUNG, M.E.J. *Radiological Physics*, 3rd Edn. Pp.788. ISBN 7186-0452-0. (H.K. Lewis & Co: 1983.) £35.

Chemistry

- CEAUSESCU, E. *Stereospecific Polymerization of Isoprene*. Pp.300. ISBN 0-08-029987-3. (Pergamon: 1983.) £30, \$60.
- CRUMP, G.B. (ed.). *Petroanalysis '81: Advances in Analytical Chemistry in the Petroleum Industry 1975-1982*. Pp.456. ISBN 0-471-26217-X. (Wiley: 1982.) £36, \$75.
- ELEY, D.D., PINES, H. and WEISZ, P.B. (eds). *Advances in Catalysis*, Vol.31. Pp.351. ISBN 0-12-007381-7. (Academic: 1982.) \$52.50.

JØRGENSEN, P. and ODDERSHEDE, J. *Problems in Quantum Chemistry*. Pp.286. ISBN 0-201-05486-8. (Addison-Wesley: 1983.) £20.20.

KURODA, P.K. *The Origin of the Chemical Elements and the Oklo Phenomenon*. Pp.165. ISBN 3-540-11679-6/0-387-11679-6. (Springer-Verlag: 1982.) DM 92, \$36.80.

LAWRENCE, J.F. (ed.). *Trace Analysis*, Vol.2. Pp.275. ISBN 0-12-682102-X. (Academic: 1982.) \$35.

MITTAL, K.L. (ed.). *Physicochemical Aspects of Polymer Surfaces*, Vol.1 and Vol.2. Pp.1250. Vol.1, ISBN 0-306-41189-X; Vol.2, ISBN 0-306-41190-3. (Plenum: 1983.) Vol.1 \$65; Vol.2 \$65.

NOZAKI, H. (ed.). *Current Trends in Organic Synthesis*. Proceedings of the 4th International Conference held August 1982, Japan. Pp.429. ISBN 0-08-029217-8. (Pergamon: 1983.) £45, \$90.

RAMASESHAN, S., RICHARDSON, M.F. and WILSON, A.J.C. (eds). *Crystallographic Statistics: Progress and Problems*. Pp.313. No ISBN. (Indian Academy of Sciences: 1982.) Np.

STONE, F.G.A. and WEST, R. (eds). *Advances in Organometallic Chemistry*, Vol.21. Pp.304. ISBN 0-12-031121-6. (Academic: 1982.) \$52.

Technology

GRANT, M.A., DONALDSON, I.G. and BIXLEY, P.F. *Geothermal Reservoir Engineering*. Pp.369. ISBN 0-12-295620-6. (Academic: 1982.) \$45.

KAZAN, B. (ed.). *Advances in Image Pickup and Display*, Vol.5. Pp.288. ISBN 0-12-022105-5. (Academic: 1982.) \$56.

LATANISION, R.M. and PICKENS, J.R. (eds). *Atomistics of Fracture*. Materials Science, Vol.5. Pp.1074. ISBN 0-306-41029-X. (Plenum: 1983.) \$115.

ROTHMAN, H., GREENSHIELDS, R. and CALLÉ, F.R. *Energy from Alcohol: The Brazilian Experience*. Pp.188. ISBN 0-8131-1479-9. (University Press of Kentucky: 1983.) \$20.

TEWARI, S.K. *Recent Trends in Wind Energy*. Pp.146. No ISBN. (Indian National Scientific Documentation Centre: 1982.) Np.

TU, K.N. and ROSENBERG, R. (eds). *Preparation and Properties of Thin Films*. Treatise on Materials Science and Technology, Vol.24. Pp.337. ISBN 0-12-341824-0. (Academic: 1982.) \$48.

Computer Science

CLARKE, M.R.B. (ed.). *Advances in Computer Chess* 3. Pp.182. ISBN 0-08-026898-6. (Pergamon: 1982.) £12.50, \$25.

COOPER, J.W. *The Minicomputer in the Laboratory: With Examples Using the PDP-11*, 2nd Edn. Pp.381. ISBN 0-471-09012-3. (Wiley: 1983.) £25.75, \$38.50.

FRUDE, N. *The Intimate Machine: Close Encounters with the New Computers*. Pp.190. Hbk ISBN 0-712-60070-1; pbk ISBN 0-712-60071-X. (Century, London: 1983.) Hbk £9.95; pbk £4.95.

Earth Sciences

BARKER, D.S. *Igneous Rocks*. Pp.417. ISBN 0-13-450692-8. (Prentice-Hall: 1983.) £29.70.

EMBRY, A.F. and BALKWILL, H.R. (eds). *Arctic Geology and Geophysics*. Proceedings of the 3rd International Symposium. Pp.552. ISBN 0-920230-19-9. (Canadian Society of Petroleum Geologists: 1982.) \$36 (Canada), \$38 (elsewhere).

KENNETT, B.L.N. *Seismic Wave Propagation in Stratified Media*. Pp.342. ISBN 0-521-23933-8. (Cambridge University Press: 1983.) £30.

SALTZMAN, B. (ed.). *Advances in Geophysics*, Vol.24. Pp.328. ISBN 0-12-018824-4. (Academic: 1982.) \$46.

WAIT, J.R. *Geo-Electromagnetism*. Pp.268. ISBN 0-12-730880-6. (Academic: 1982.) \$34.

WHITMORE, F.C. Jr. and WILLIAMS, M.E. (eds). *Resources for the Twenty-First Century*. Geological Survey Professional Paper 1193. Proceedings of the International Centennial Symposium held October 1979, Virginia. Pp.345. No ISBN. (U.S. Government Printing Office) Np.

Environmental Sciences

CHAMP, M.A. and PARK, P.K. *Global Marine Pollution Bibliography: Ocean Dumping of Municipal and Industrial Wastes*. Pp.399. ISBN 0-306-65205-6. (IFI/Plenum: 1982.) Np.

COPE, C.B., FULLER, W.H. and WILLETTS, S.L. *The Scientific Management of Hazardous Wastes*. Pp.480. ISBN 0-521-25100-1. (Cambridge University Press: 1983.) £35.

Biological Sciences

LEIBLICH, I. (ed.). *Genetics of the Brain*. Pp.491. ISBN 0-444-80438-2. (Elsevier Biomedical/North Holland: 1982.) Dfl. 330, \$140.50.

LLOYD, C.W. (ed.). *The Cytoskeleton in Plant Growth and Development*. Pp.450. ISBN 0-12-453780-4. (Academic: 1983.) £32.40, \$60.

LOCQUIN, M. and LANGERON, M. *Handbook of Microscopy*. Pp.322. ISBN 0-408-10679-4. (Butterworth: 1983.) £50.

LUBEC, G. (ed.). *Renal Immunology*. Contributions to Nephrology, Vol. 35. Pp.194. Pbk ISBN 3-8055-3587-2. (Karger: 1983.) SwFr.97, DM116, \$58.25.

PATIN, S.A. *Pollution and the Biological Resources of the Oceans*. Pp.287. ISBN 0-408-10840-1. (Butterworth Scientific: 1982.) £38.50.

PECK, W.A. (ed.). *Bone and Mineral Research, Annual: A Yearly Survey of Developments in the Field of Bone and Mineral Metabolism*. Pp.377. ISBN 90-219-3065-X/0-444-90239-2. (Excerpta Medica/Elsevier Biomedical/North Holland: 1983.) \$80.75, Dfl.190.

PIMM, S.L. *Food Webs*. Pp.219. Hbk ISBN 0-412-23100-X; pbk ISBN 0-412-23110-7. (Chapman & Hall: 1982.) Hbk £15; pbk £7.50.

SARKANEN, K.V., TILLMAN, D.A. and JAHN, E.C. (eds). *Progress in Biomass Conversion*, Vol. 3. Pp.291. ISBN 0-12-535903-9. (Academic: 1982.) \$24.50.

SEPKOSKI, J.J. Jr. *Contributions in Biology and Geology*, No. 51. Pp. 125. Pbk ISBN 0-893260-081-9. (Milwaukee Public Museum Press: 1982.) \$15.

SPEAR, N.E. and MILLER, R.R. (eds). *Information Processing in Animals: Memory Mechanisms*. Pp.390. ISBN 0-89859-157-0. (Lawrence Erlbaum Associates: 1981.) £24.95.

STEINER, W.W.M. *et al.* (eds). *Recent Developments in the Genetics of Insect Disease Vectors*. Pp.665. ISBN 0-87563-224-6. (Stipes: 1982.) \$26.

TAYLOR, D.A. *Mind: A Scientist's View of How the Mind Works — and How To Make It Work For You*. Pp.239. Hbk ISBN 0-712-60136-8; pbk ISBN 0-712-60137-6. (Century, London: 1983.) Hbk £9.95; pbk £4.95.

THORBECKE, G.J. and LESLIE, G.A. (eds). *Immunoglobulin D: Structure and Function*. Annals of the New York Academy of Sciences, Vol. 399. Pp. 410. Hbk ISBN 0-89766-188-5; pbk ISBN 0-89766-189-3. (New York Academy of Sciences: 1982.) Hbk np; pbk \$80.

TIPSON, R.S. and HORTON, D. (eds). *Advances in Carbohydrate Chemistry and Biochemistry*, Vol. 40. Pp.452. ISBN 0-12-007240-8. (Academic: 1982.) \$65.

TUDOR, E.R. and TUDOR, E.M. *Understanding the Human Body*. Pp.431. Pbk ISBN 0-85896-819-3/0-273-01952-X. (Pitman/Richard Clay (SE Asia): 1983.) £7.50.

WEBB, D.A. and SCANNEL, M.J.P. *Flora of Connemara and The Burren*. Pp.322. ISBN 0-521-23395-X. (Cambridge University Press: 1983.) £35.

WEST, N.E. (ed.). *Temperate Deserts and Semi-Deserts. Ecosystems of the World*, Vol. 5. Pp. 522. ISBN 0-444-41931-4. (Elsevier Biomedical/North Holland: 1983.) \$170.25, Dfl.400.

WESTBROEK, P. and DE JONG, E.W. (eds). *Bio-mineralization and Biological Metal Accumulation: Biological and Geological Perspectives*. Papers presented at the 4th International Symposium held June 1982, The Netherlands. Pp. 533. ISBN 90-277-1515-7. (Reidel: 1983.) Dfl. 160, \$69.50.

ZBINDEN, G. *et al.* (eds). *Application of Behavioral Pharmacology in Toxicology*. Pp.409. ISBN 0-89004-902-5. (Raven: 1982.) \$48.98.

BOOKS RECEIVED

ZIEGLER, B. Introduction to Palaeobiology: General Palaeontology. Pp.225. Hbk ISBN 0-85312-211-3/0-470-27552-9; pbk ISBN 0-85312-531-7. (Ellis Horwood/Halsted: 1983.) Hbk £30, \$60.30; pbk np.

Applied Biological Sciences

ARCOS, J.C., WOO, Y-T. and ARGUS, M.F. Chemical Induction of Cancer: Structural Bases and Biological Mechanisms, Vol.3A. Pp.780. ISBN 0-12-059303-3. (Academic: 1982.) \$74.50.

BASERGA, R. (ed.). Cell Proliferation, Cancer, and Cancer Therapy. Annals of the New York Academy of Sciences, Vol.397. Pp.328. Hbk ISBN 0-89766-184-2; pbk ISBN 0-89766-185-0. (New York Academy of Sciences: 1982.) Hbk np; pbk \$65.

BULBROOK, R.D. and TAYLOR, D.J. (eds). Commentaries on Research in Breast Disease, Vol.2. Pp.180. ISBN 0-8451-1901-X. (Alan R. Liss: 1981.) £22.

COLLINS, J.A., MURAWSKI, K. and SHATER, A.W. (eds). Massive Transfusion in Surgery and Trauma. Proceedings of the XIVth Annual Scientific Symposium of the American Red Cross held May 1982, Washington, DC. Progress in Clinical and Biological Research, Vol.108. Pp.328. ISBN 0-8451-0108-0. (Alan R. Liss: 1982.) £25.

CUELLO, A.C. (ed.). Brain Microdissection Techniques. Pp.186. Hbk ISBN 0-471-10523-6; pbk ISBN 0-471-90019-2. (Wiley: 1983.) Hbk np; pbk £11.50, \$23.

CUMMINGS, N.B., MICHAEL, A.F. and WILSON, C.B. (eds). Immune Mechanisms in Renal Disease. Pp.564. ISBN 0-306-40948-8. (Plenum: 1983.) Np.

Drought Resistance in Crops with Emphasis on Rice. Pp.414. Pbk ISBN 971-104-078-6. (International Rice Research Institute: 1982.) Np.

DUNN, M.J., PARTONO, C. and CINOTTI, G.A. (eds). Prostaglandins and the Kidney: Biochemistry, Physiology, Pharmacology, and Clinical Applications. Pp.414. ISBN 0-306-41054-0. (Plenum: 1983.) Np.

GLASS, R.H. and ERICSSON, R.J. Getting Pregnant in the 1980s: New Advances in Infertility Treatment and Sex Preselection. Pp.164. ISBN 0-520-04828-8. (University of California Press: 1983.) £8.75.

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Supplement 4: Chemicals, Industrial Processes and Industries Associated with Cancer in Humans — IARC Monographs. Vol. 1-29. Pp.467. Pbk ISBN 92-832-1406-4. (WHO: 1982.) SwFr. 60, \$30.

ITO, Z., KUTSUZAWA, T. and YASUI, N. (eds). Cerebral Ischemia — An Update — Proceedings of the 3rd AKITA International Joint Conference. Pp.273. ISBN 90-219-0599-X/0-444-90290-2. (Excerpta Medica/Elsevier Biomedical/North Holland: 1983.) \$63.75, Dfl.150.

JELJASZEWICZ, J., PULVERER, G. and ROSKOWSKI, W. (eds). Bacteria and Cancer. Pp.464. ISBN 0-12-383820-7. (Academic: 1983.) £21.40, \$39.50.

KESSEL, D. and DOUGHERTY, T.J. (eds). Porphyrin Photosensitization. Advances in Experimental Medicine and Biology, Vol.160. Pp.294. ISBN 0-306-41193-8. (Plenum: 1983.) Np.

KISSIN, B. and BEGLEITER, H. (eds). The Pathogenesis of Alcoholism: Psychosocial Factors. The Biology of Alcoholism, Vol.7. Pp.695. ISBN 0-306-41052-4. (Plenum: 1983.) Np.

LANGENBACH, R., NESNOW, S. and RICE, J.M. (eds). Organ and Species Specificity in Chemical Carcinogenesis. Basic Life Sciences, Vol.24. Pp.693. ISBN 0-306-41184-9. (Plenum: 1983.) Np.

LARON, Z. and GLATZTER, A. (eds). Recent Progress in Medico-Social Problems in Juvenile Diabetics. Pediatric and Adolescent Endocrinology, Vol.11. Proceedings of the 5th International Beilinson Symposium held October 1981, Herzliya-on-Sea. Pp.210. ISBN 3-8055-3594-5. (Karger: 1983.) SwFr.153, DM 183, \$91.75.

LEVIN, M.A. *et al.* Applied Genetic Engineering: Future Trends and Problems. Pp.346. ISBN 0-8155-0925-1. (Noyes: 1983.) \$24.

LIFSHITZ, F. (ed.). Pediatric Nutrition: Infant Feedings—Deficiencies—Diseases. Clinical Disorders in Pediatric Nutrition Series, Vol.2. Pp.648. ISBN 0-8247-1430-X. (Marcel Dekker: 1982.) SwFr.156.

LOWENTHAL, A., MORI, A. and MARESCAU, B. (eds). Urea Cycle Diseases. Advances in Experimental Medicine and Biology, Vol.153. Pp.528. ISBN 0-306-41037-0. (Plenum: 1982.) Np.

MALICK, J.B., ENNA, S.J. and YAMAMURA, H.I. (eds). Anxiolytics: Neurochemical, Behavioral, and Clinical Perspectives. Pp. 232. ISBN 0-89004-731-6. (Raven: 1982.) \$48.36.

MARCELLE, R., CLIJSTERS, H. and VAN POUCKE, M. (eds). Effects of Stress on Photosynthesis. Advances in Agricultural Biotechnology. Proceedings of a Conference held August 1982, Belgium. Pp. 388. ISBN 90-247-2799-5. (Martinus Nijhoff: 1983.) Dfl. 120, \$52.

MUGGIA, F.M. (ed.). Cancer Chemotherapy, Vol. 1. Cancer Treatment and Research, Vol. 7. Pp. 346. ISBN 90-247-2713-8. (Martinus Nijhoff: 1983.) Dfl. 160, \$69.50.

PAPADATOS, C.J. and BARTSOCAS, C.S. (eds). Skeletal Dysplasias. Proceedings of the 3rd International Seminar held May 1982, Greece. Progress in Clinical and Biological Research, Vol. 104. Pp. 572. ISBN 0-8451-0104-8. (Alan R. Liss: 1982.) £46.

PRASANSUK, S., BUNNAG, C. and SRISOMBOON, P. (eds). Aspects of Modern Otolaryngological Practice. Advances in Oto-Rhino-Laryngology, Vol. 29. Pp. 236. ISBN 3-8055-3592-9. (Karger: 1983.) SwFr. 164, DM 196, \$98.25.

SIROIS, P. and ROLA-PLESZCZYNSKI, M. (eds). Immunopharmacology. Research Monographs in Immunology, Vol. 4. Pp. 426. ISBN 0-444-80416-1. (Elsevier Biomedical/North Holland: 1982.) Dfl. 230, \$97.75.

STEWART, D.A. (ed.). Children With Sex Chromosome Aneuploidy: Follow-up Studies. Birth Defects: Original Article Series, Vol. 18, No. 4. Pp. 276. ISBN 0-8451-1052-7. (Alan R. Liss: 1982.) £41.

STUHRMANN, H.B. (ed.). Uses of Synchrotron Radiation in Biology. Pp. 348. ISBN 0-12-674850-0. (Academic: 1982.) £36.40, \$67.50.

US DEPARTMENT OF HEALTH AND HUMAN SERVICES. Research Frontiers in Aging and Cancer: International Symposium for the 1980's. National Cancer Institute Monograph 60. Pp. 302. (NIH: 1982.) \$15 (U.S., Canada, Mexico), \$18.75 (elsewhere).

WALLACE, J.E. *et al.* Review Questions in Analytical Toxicology. Pp. 184. Pbk ISBN 0-931890-09-8. (Biomedical Publications: 1982.) £14, \$22.

WISEMAN, A. (ed.). Topics in Enzyme and Fermentation Biotechnology, Vol. 7. Pp. 314. ISBN 0-85312-465-5/0-470-27366-6. (Ellis Horwood/Halsted: 1983.) £32.50, \$65.35.

Sociology

BREUER, G. Sociobiology and the Human Dimension. Pp. 286. Hbk ISBN 0-521-24544-3; pbk ISBN 0-521-28778-2. (Cambridge University Press: 1983.) Hbk £22.50; pbk £8.95.

MRAZEK, P.B. and KEMPE, C.H. (eds). Sexually Abused Children and their Families. Pp. 271. Hbk ISBN 0-08-026796-3; pbk ISBN 0-08-030194-0. (Pergamon: 1982.) Hbk np; pbk £6.50, \$13.

WARWICK, D.P. Bitter Pills: Population Policies and Their Implementation in Eight Developing Countries. Pp. 229. ISBN 0-521-24347-5. (Cambridge University Press: 1982.) £17.50.

WEIDLICH, W. and HAAG, G. Concepts and Models of a Quantitative Sociology: The Dynamics of Interacting Populations. Springer Series in Synergetics, Vol. 14. Pp. 217. ISBN 3-540-11358-4/0-387-11358-4. (Springer-Verlag: 1983.) DM 78, \$31.20.

Psychology

DAVIDSON, R.J., SCHWARTZ, G.E. and SHAPIRO, D. (eds). Consciousness and Self-Regulation: Advances in Research and Theory, Vol. 3. Pp. 221. ISBN 0-306-41214-4. (Plenum: 1983.) Np.

GRAHAME-SMITH, D.G. (ed.). Part 1: Preclinical Psychopharmacology. Pp. 456. ISBN 90-219-3066-8/0-444-90240-6. (Excerpta Medica/Elsevier Biomedical/North Holland: 1983.) \$80.75, Dfl. 190. Set £153.25, Dfl. 360.

HIPPIUS, H. and WINOKUR, G. (eds). Part 2: Clinical Psychopharmacology. Pp. 482. ISBN 90-219-3078-1/0-444-90282-1. (Excerpta Medica/Elsevier Biomedical/North Holland: 1983.) \$85, Dfl. 200.

WING, J.K. and WING, L. (eds). Psychoses of Uncertain Aetiology. Handbook of Psychiatry, Vol. 3. Pp. 313. Hbk ISBN 0-521-24101-4; pbk ISBN 0-521-28438-4. (Cambridge University press: 1983.) Hbk £27.50; pbk £9.95.

General

BRADLEY MARTIN, E. and BRADLEY MARTIN, E. Run Rhino Run. Pp.136. ISBN 0-7011-2632-9. (Chatto & Windus: 1982.) £9.95.

FULLER, R.B. in collaboration with APPLEWHITE, E.J. Synergetics: Explorations in the Geometry of Thinking. Pp.876. Pbk ISBN 0-02-065320-4. (Collier Macmillan, Middlesex: 1982.) £9.95.

GEORGE, S. and PAIGE, N. Food for Beginners. Pp. 173. Hbk ISBN 0-906495-85-7; pbk ISBN 0-906495-84-9. (Writers and Readers: 1982.) Hbk np; pbk £2.50 (UK), \$5.95 (Austr.), \$5.95 (Canada).

GORBATY, M.L., LARSEN, J.W. and WENDER, I. (eds). Coal Science, Vol. 1. Pp. 293. ISBN 0-12-150701-7. (Academic: 1982.) \$32.50.

GULATI, L. Profiles in Female Poverty: A Study of Five Poor Working Women in Kerala. Pp.179. ISBN 0-08-027922-8. (Pergamon: 1982.) £10, \$23.

GUNTHEROTH, W.G. Crib Death: The Sudden Infant Death Syndrome. Pp.224. ISBN 087-993-175-2. (Futura: 1982.) \$60.75.

HALSTEAD, B. A Brontosaur: The Life Story Unearthed. Pp.32. ISBN 0-00-104111-8. (Collins: 1982.) £3.95.

HAYNES, R. (ed.). Environmental Science Methods. Pp.404. Hbk ISBN 0-412-23280-4; pbk ISBN 0-412-23290-1. (Chapman & Hall: 1982.) Hbk np; pbk £9.95.

LEVY, M., BASDEVANT, J.-L. *et al.* (eds). Fundamental Interactions: Cargèse 1981. NATO Advanced Study Institutes Series, Vol.85. Pp.696. ISBN 0-306-41116-4. (Plenum: 1982.) \$89.50.

MANSFIELD, S. The Gestalts of War: An Inquiry Into Its Origins and Meanings as a Social Institution. Pp.274. ISBN 0-385-27219-7. (The Dial Press: 1982.) \$16.95.

MCNEILL, W.H. The Pursuit of Power: Technology, Armed Force, and Society Since AD 1000. Pp.405. ISBN 0-631-13134-5. (Basil Blackwell: 1983.) £15.

NEVATIA, R. Machine Perception. Pp. 209. ISBN 0-13-541904-2. (Prentice-Hall: 1982.) £22.40.

SMITH, D.K. Network Optimisation Practice: A Computational Guide. Pp.237. UK ISBN 0-85312-403-5; pbk ISBN 0-85312-445-0; US ISBN 0-470-27347-X. (Ellis Horwood/Halsted: 1982.) Hbk £22.50, \$45.25; pbk np.

SPETGANG, T. and WELLS, M. The Children's Solar Energy Book. Pp.156. Hbk ISBN 0-8069-3119-1; pbk ISBN 0-8069-7584-9. (Sterling: 1982.) Hbk £6.95; pbk np.

STANNARD, R. Science and the Renewal of Belief. Pp.213. Pbk ISBN 334-01455-7. (SCM Press: 1982.) £2.95.

TAUBE, M. Evolution of Matter and Energy: In Cosmic and Planetary Scale. No ISBN. (Hausdruckerei des Eidg. Instituts für Reaktorforschung, Switzerland: 1982.) Np.

TEICHLER-ZALLEN, D. and CLEMENTS, C.D. (eds). Science and Morality: New Directions in Bioethics. Pp.291. ISBN 0-669-04406-7. (Lexington: 1982.) £19.50.

WATSON, P. Sea Shepherd: My Flight for Whales and Seals. Pp.258. ISBN 0-393-01499-1. (W.W. Norton & Co: 1982.) £10.50, \$14.95.

WOOD, M.A. (ed.). Guidelines for a Voluntary Cancer Organization. UICC Technical Report Series, Vol.65. Pp.176. Pbk ISBN 92-9018-065-X. (International Union Against Cancer/Hans Huber: 1982.) SwFr. 28, \$14.

WOODEN, W.S. and PARKER, J. Men Behind Bars: Sexual Exploitation in Prison. Pp.264. ISBN 0-306-41074-5. (Plenum: 1982.) \$15.95.

ZOHARY, M. Plants of the Bible. Pp.223. ISBN 0-521-24926-0. (Cambridge University Press: 1982.) £9.50.